Algèbre linéaire 1

R. Cairoli
Professeur à l'Ecole polytechnique
fédérale de Lausanne

Presses polytechniques romandes

Edités par les Presses polytechniques romandes
(documentation détaillée sur demande aux
Presses polytechniques romandes,
EPFL-Centre Midi, CH-1015 Lausanne, Suisse).

Calcul différentiel et intégral
Jacques Douchet et Bruno Zwahlen

1. Fonctions réelles d'une variable réelle

2. Fonctions réelles de plusieurs variables réelles

3. Exercices
(parution 1987)

Analyse non standard
Alain Robert

Algèbre linéaire, tomes 1 et 2
R. Cairoli

Initiation aux probabilités
Sheldon M. Ross

Si vous désirez être tenu au courant des publications de l'éditeur de cet ouvrage, envoyez vos nom, prénom et adresse aux Presses polytechniques romandes (EPFL-Centre Midi, CH-1015 Lausanne, Suisse) qui vous enverront leur catalogue général.

ISBN 2-88074-110-6
Composition et impression Schüler SA, 2500 Bienne

Imprimé en Suisse

Préface

L'enseignement de l'algèbre linéaire s'est considérablement développé au cours des deux dernières décennies. De nos jours, presque toutes les sections d'études scientifiques et techniques, et notamment les sections d'ingénieurs, incluent l'algèbre linéaire dans la formation de base de leurs étudiants. Un ouvrage qui s'ajoute aux nombreux déjà existants dans la littérature peut donc encore se justifier et prétendre répondre à des exigences restées insatisfaites.

A l'exception de quelques-uns, les sujets développés dans ce livre sont classiques et servent normalement de base à la réalisation de tout livre d'algèbre linéaire de même niveau. Ce qui distingue celui-ci des autres réside dans l'arrangement de la matière et surtout dans sa présentation, qui emprunte beaucoup à la géométrie ordinaire et vise à développer chez le lecteur une compréhension intuitive. Un autre élément distinctif est certainement la place accordée à la notion d'espace affine et à l'étude de la géométrie affine à plusieurs dimensions.

Dans l'élaboration de la matière, l'auteur s'est prévalu de l'expérience de plusieurs années d'enseignement à des sections d'ingénieurs de l'Ecole polytechnique fédérale de Lausanne. C'est à travers cette expérience que l'exposition purement formelle conçue initialement a cédé la place à un discours plus parlé et, oserait-on dire, plus humain.

Conscient du fait que les concepts abstraits ne deviennent clairs qu'à travers les exemples, l'auteur s'est constamment soucié de favoriser la compréhension par des motivations, des commentaires et des illustrations.

Outre créer les conditions propices à une meilleure assimilation des théorèmes et des techniques de calcul de l'algèbre linéaire, cet ouvrage veut inciter le lecteur à un travail de recherche personnel. Quelques-uns des nombreux exercices placés à la fin de chaque chapitre ont pour but d'encourager une telle activité.

Ce livre a été écrit à l'intention des étudiants du premier cycle d'études des écoles d'ingénieurs de niveau universitaire, mais il s'adresse également aux étudiants en mathématique et en physique orientés vers les applications. Il peut en outre venir en aide aux scientifiques à la recherche de méthodes algébriques leur permettant d'apporter des éléments de réponse aux problèmes qu'ils rencontrent, ainsi qu'aux maîtres du degré secondaire désireux de savoir vers quels programmes conduit leur enseignement.

Remerciements

Je remercie vivement Monsieur Jean-Claude Evard de l'intérêt constant qu'il a apporté à la réalisation de cet ouvrage. Ses critiques et ses suggestions m'ont permis d'améliorer la présentation de nombreux points délicats du texte.

Je remercie Monsieur Laurent Perruchoud d'avoir lu plusieurs parties du manuscrit et de m'avoir signalé quelques imprécisions.

Je remercie également Monsieur Claude El-Hayek d'avoir contrôlé la version définitive du manuscrit, Monsieur Klaus-Dieter Semmler d'avoir réalisé les figures représentant les quadriques, Monsieur Jean-François Casteu d'avoir effectué les dessins et Madame Pascale Deppierraz pour la compétence avec laquelle elle s'est occupée des problèmes d'édition.

Conventions

1. Découpage du texte

Ce livre est divisé en deux tomes. Chaque tome se compose de cinq chapitres numérotés respectivement de 1 à 5 et de 6 à 10. Le second tome contient en outre un appendice repéré par la lettre A. Chaque chapitre est divisé en sections et chaque section en paragraphes. Les sections sont repérées par une double numérotation et les paragraphes par une triple numérotation. Par exemple, 7.2 renvoie à la deuxième section du septième chapitre et 7.2.4 au quatrième paragraphe de cette section. La dernière section de chaque chapitre rassemble les exercices sur la matière traitée dans le chapitre. Ces exercices sont numérotés de la même façon que les paragraphes. Ainsi, 7.4.11 désigne le onzième exercice de la quatrième section du septième chapitre. Les figures sont repérées par l'abréviation Fig. suivie de deux nombres, le premier indiquant le chapitre et le deuxième la figure. Par exemple, Fig. 7.3 désigne la troisième figure du septième chapitre.

2. Conventions sur les nombres

Les lettres $\mathbb{R}$ et $\mathbb{C}$ désigneront respectivement le corps des nombres réels et le corps des nombres complexes. Tout au long des dix chapitres de ce livre, le terme «nombre» aura le sens de «nombre réel». Dans l'appendice consacré à l'extension de certains résultats aux nombres complexes, il sera précisé dans quelles circonstances le terme «nombre» prendra la signification de «nombre complexe».

Les nombres seront généralement désignés par des lettres grecques minuscules telles que α, β, γ, λ, μ, ν, ou par des lettres latines minuscules telles que a, b, c, d. Les nombres entiers seront appelés plus simplement entiers. Ils seront désignés par des lettres latines minuscules telles que n, k, i, j, l, m.

Nous dirons qu'un nombre est *positif* (*négatif*) s'il est supérieur (inférieur) à zéro. Nous dirons qu'il est *non négatif* s'il est supérieur ou égal à zéro.

3. Avertissement concernant l'emploi des adjectifs numéraux

Tout au long de ce livre, par deux, trois, ..., n objets, nous entendrons (sauf mention explicite du contraire) deux, trois, ..., n objets distincts.

4. Familles d'éléments d'un ensemble

Dans ce livre, nous appellerons *famille finie* d'éléments d'un ensemble E tout n-uplet (ordonné) d'éléments de E, c'est-à-dire tout élément du produit cartésien

$(x_1, x_2, ..., x_n)$ et dirons que x_i est le *i-ième terme* ou le *terme d'indice i*. Nous dirons en outre que l'ensemble $\{1, 2, ..., n\}$ est l'*ensemble d'indices*.

On remarquera que notre définition n'exclut pas que $x_i = x_j$ pour des indices i et j différents, ni même que $x_i = x$ pour tout indice i, x désignant un élément de E. On remarquera encore que $(x_1, x_2, ..., x_n) = (x'_1, x'_2, ..., x'_n)$ si et seulement si $x_1 = x'_1, x_2 = x'_2, ..., x_n = x'_n$.

Nous appellerons les familles à deux et à trois termes respectivement *couples* et *triplets*.

Les *familles infinies* (ou *suites*) sont définies similairement en prenant comme ensemble d'indices l'ensemble des entiers positifs. Elles seront désignées par le symbole $(x_1, x_2, ...)$. Précisons toutefois qu'elles apparaîtront très rarement dans ce livre.

Ajoutons quelques lignes de commentaire à la notion de famille finie. Dans l'étude de nombreuses questions, telles l'indépendance linéaire ou la génération de sous-espaces ou l'orthogonalité, l'incorporation d'un ordre à la définition de famille finie est inutile. A ce type de questions, les familles «non ordonnées», c'est-à-dire les applications d'un ensemble fini I dans E (notées habituellement par $(x_i)_{i \in I}$) conviennent parfaitement. Par contre, dans d'autres contextes, notamment dans certaines relations avec les matrices (en raison de leur représentation sous la forme de tableaux) ou dans des questions concernant l'orientation, l'ordre de disposition des termes de la famille joue un rôle important. Dans ce livre, par souci de simplicité, nous n'avons pas jugé nécessaire d'utiliser deux notions de famille finie. Le lecteur intéressé pourra facilement déceler de lui-même les parties qui s'énoncent plus naturellement en termes de familles finies «non ordonnées».

Table des matières

TOME 1

TOME 2

Espaces vectoriels et espaces affines

1.1 UN MODÈLE D'ESPACE VECTORIEL

1.1.1 Introduction

Des notions physiques telles que la force ou la vitesse sont caractérisées par une direction, un sens et une intensité. Ce triple caractère est mis en évidence par les flèches. Celles-ci sont à l'origine de la notion de vecteur et en constituent l'exemple le plus suggestif. Bien que leur nature soit essentiellement géométrique, c'est leur aptitude à se lier les unes aux autres, donc leur comportement algébrique, qui retiendra principalement notre attention. Partagé en classes d'équivalence et muni de deux opérations appelées addition et multiplication par un scalaire, l'ensemble qu'elles forment représente le modèle classique d'un espace vectoriel. Un de nos premiers objectifs est la description détaillée de ce modèle.

1.1.2 Notion de flèche

Nous désignerons par $\mathscr{G}$ l'espace ordinaire de la géométrie élémentaire et par P, Q, ... ses points. Nous appellerons *flèche* tout segment de droite orienté. La flèche d'origine P et d'extrémité Q sera notée PQ (fig. 1.1). Il est évident que toute flèche est caractérisée par sa direction, son sens, son intensité ou grandeur et son origine.

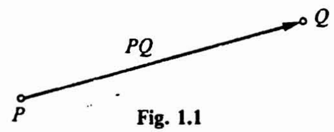

Fig. 1.1

1.1.3 Ensemble des vecteurs

Nous dirons que deux flèches sont équivalentes si elles ont la même direction, le même sens et la même intensité. Partageons l'ensemble des flèches en classes d'équivalence: deux flèches appartiennent à une même classe si et seulement si elles sont équivalentes. Nous dirons que chacune de ces classes est un *vecteur*. Rangeons, en outre, les flèches dégénérées (c'est-à-dire de la forme PP) en une classe distinguée que nous appellerons *vecteur nul* et noterons $\mathbf{0}$. L'ensemble des vecteurs

ainsi définis sera désigné par V. Il faut souligner que les éléments de V sont des classes de flèches et non pas des flèches individuelles. Il est cependant clair qu'une flèche quelconque suffit à déterminer la classe à laquelle elle appartient et il est donc naturel de l'appeler *représentant* de la classe ou du vecteur.

Dans ce livre, les vecteurs seront désignés par des lettres latines minuscules imprimées en caractère gras ou par des couples de lettres latines majuscules surmontées d'une flèche (par exemple, $\overrightarrow{PQ}$ désigne le vecteur déterminé par la flèche PQ). Dans les figures, les flèches seront toutefois désignées par le symbole du vecteur qu'elles représentent.

1.1.4 Addition de vecteurs

Traçons le représentant d'un vecteur $\mathbf{y}$ à partir de l'extrémité d'un représentant d'un vecteur $\mathbf{x}$. La flèche dont l'origine est celle du représentant de $\mathbf{x}$ et l'extrémité celle du représentant de $\mathbf{y}$ détermine un vecteur que nous noterons $\mathbf{x} + \mathbf{y}$ et appelerons *somme* de $\mathbf{x}$ et $\mathbf{y}$. L'opération qui associe à tout couple de vecteurs leur somme s'appelle *addition vectorielle* (fig. 1.2).

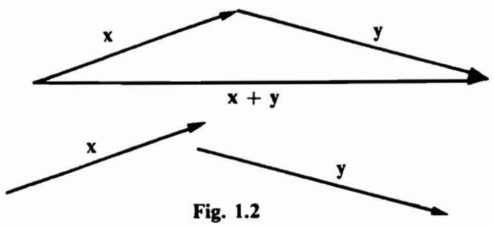

Fig. 1.2

A l'aide d'une figure, il est facile de montrer que l'opération d'addition vectorielle est *associative* et *commutative*, autrement dit, que

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$$

et

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}.$$

Il est en outre évident que le vecteur nul $\mathbf{0}$ est l'élément neutre de l'addition vectorielle, autrement dit, que

$$\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x},$$

et que

$$\mathbf{x} + (-\mathbf{x}) = \mathbf{0},$$

où $-\mathbf{x}$ désigne le *vecteur opposé* de $\mathbf{x}$, c'est-à-dire le vecteur dont les représentants ont la même direction et la même intensité que ceux de $\mathbf{x}$, mais le sens opposé (fig. 1.3).

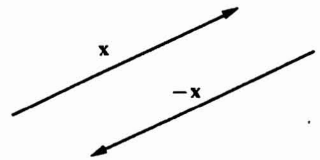

Fig. 1.3

1.1.5 Soustraction de vecteurs

L'opération inverse de l'addition vectorielle est la ***soustraction** vectorielle*. Soustraire un vecteur revient à additionner le vecteur opposé (fig. 1.4):

$$\mathbf{x} - \mathbf{y} = \mathbf{x} + (-\mathbf{y}).$$

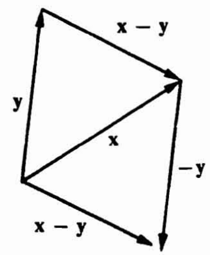

Fig. 1.4

1.1.6 Remarque

L'addition s'étend, par récurrence, au cas d'une famille finie quelconque de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$:

$$((\mathbf{x}_1 + \mathbf{x}_2) + \mathbf{x}_3) + \ldots .$$

En vertu de l'associativité, ces additions successives peuvent être effectuées dans n'importe quel ordre, ce qui justifie l'écriture sans parenthèses

$$\mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_k.$$

1.1.7 Multiplication par un scalaire

Dans ce livre «scalaire» sera synonyme de «nombre». Rappelons en outre que, sauf indication contraire, «nombre» signifie «nombre réel». Le vecteur $\alpha\mathbf{x}$, appelé *produit du nombre α par* $\mathbf{x}$, est défini de la manière suivante: prenons une flèche représentative de $\mathbf{x}$ et construisons une flèche de même direction, de même sens ou de sens opposé, suivant que α est positif ou négatif, et d'intensité $|\alpha|$ fois l'intensité de la flèche initiale; la flèche ainsi obtenue est un représentant du vecteur $\alpha\mathbf{x}$; si $\alpha = 0$ ou $\mathbf{x} = \mathbf{0}$, nous posons $\alpha\mathbf{x} = \mathbf{0}$. L'opération qui consiste à effectuer

(7) $(-\alpha)\mathbf{x} = -(\alpha\mathbf{x})$, donc en particulier $(-1)\mathbf{x} = -\mathbf{x}$. En effet, l'opposé d'un vecteur étant unique d'après (3), il suffit de montrer que $\alpha\mathbf{x} + (-\alpha)\mathbf{x} = \mathbf{0}$. Or, par la condition (f), $\alpha\mathbf{x} + (-\alpha)\mathbf{x} = (\alpha - \alpha)\mathbf{x} = 0\mathbf{x}$ et, par (5), $0\mathbf{x} = \mathbf{0}$.

1.3 EXEMPLES D'ESPACES VECTORIELS

1.3.1 Espaces vectoriels géométriques

L'espace vectoriel géométrique V étudié dans la section 1.1 est un premier exemple d'espace vectoriel selon la définition 1.2.2. Un deuxième exemple est le plan vectoriel géométrique, c'est-à-dire l'ensemble des classes de flèches équivalentes du plan usuel de la géométrie élémentaire, muni des deux opérations introduites dans 1.1.4 et 1.1.7. Nous le désignerons également par V, mais s'il faut le distinguer du premier, nous utiliserons les symboles V^3 pour l'espace et V^2 pour le plan.

Inclure ✱

1.3.2 Vectorialisé de $\mathscr{G}$.

Soit O un point arbitrairement choisi et fixé de l'espace ponctuel $\mathscr{G}$ introduit dans 1.1.2. Définissons l'opération d'addition des points P et Q de $\mathscr{G}$ par la règle du parallélogramme: $P + Q$ est le sommet opposé à O du parallélogramme construit sur O, P, Q (fig. 1.6). Cette opération peut également être définie à l'aide des flèches: $P + Q$ est l'extrémité de la flèche d'origine P équivalente à la flèche OQ. Définissons similairement la multiplication de P par un nombre α (fig. 1.7). Muni de ces deux opérations, $\mathscr{G}$ devient un espace vectoriel appelé *vectorialisé* de $\mathscr{G}$ relativement à O. Nous désignerons cet espace par $\mathscr{G}_O$ et appellerons le point O *origine*. En identifiant chaque point P à la flèche OP, nous pouvons considérer $\mathscr{G}_O$ comme étant formé des flèches d'origine commune O.

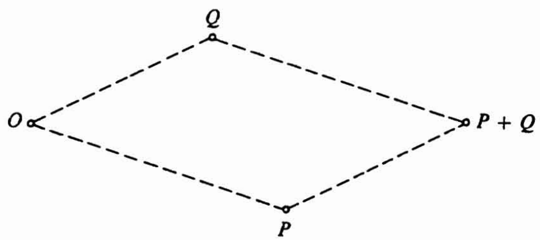

Fig. 1.6

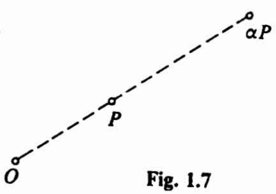

Fig. 1.7

On remarquera que si O' est une deuxième origine, $\mathscr{G}_{O'}$ et $\mathscr{G}_O$ sont égaux en tant qu'ensembles, mais différents en tant qu'espaces vectoriels (fig. 1.8).

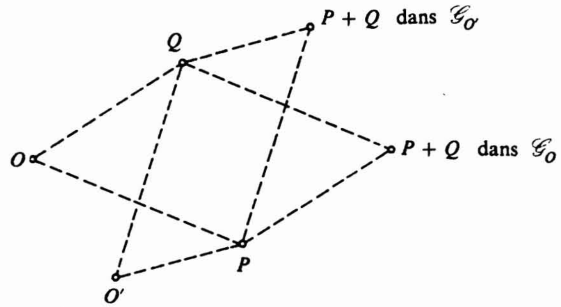

Fig. 1.8

1.3.3 Espaces $\mathbb{R}^n$

Pour tout entier positif n, $\mathbb{R}^n$ désignera l'ensemble des n-uplets de nombres disposés en colonne:

$$\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}.$$

Munissons $\mathbb{R}^n$ des opérations d'addition et de multiplication par un scalaire définies au moyen des formules suivantes:

$$\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n + b_n \end{pmatrix}, \quad \alpha \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha a_n \end{pmatrix}. \tag{1.1}$$

Ces deux opérations satisfont, de toute évidence, aux conditions (a)–(h) de la définition 1.2.2 et confèrent donc à $\mathbb{R}^n$ une structure d'espace vectoriel. Les vecteurs de cet espace seront appelés *vecteurs-colonnes*. Ils seront souvent désignés plus brièvement par $(a_i)_{1 \leqslant i \leqslant n}$ ou simplement par (a_i). Le nombre a_i sera appelé *terme d'indice i* de (a_i).

On remarquera que le vecteur nul de $\mathbb{R}^n$ est le n-uplet

$$\begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

et que

$$\begin{pmatrix} -a_1 \\ -a_2 \\ \cdot \\ \cdot \\ \cdot \\ -a_n \end{pmatrix} \text{ est l'opposé de } \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}.$$

$\mathbb{R}^1$ sera identifié à $\mathbb{R}$.

1.3.4 Un espace vectoriel fonctionnel

Soit $C_{[a,b]}$ l'ensemble des fonctions réelles continues définies dans l'intervalle fermé $[a, b]$. Nous désignerons les éléments de cet ensemble par les lettres $\mathbf{f}$, $\mathbf{g}$, ... La valeur de $\mathbf{f}$ au point t sera notée $\mathbf{f}(t)$. Dire que $\mathbf{f} = \mathbf{g}$ équivaudra donc à dire que $\mathbf{f}(t) = \mathbf{g}(t)$ pour tout t de l'intervalle $[a, b]$. De manière abrégée, nous écrirons $\mathbf{f}(t) \equiv \mathbf{g}(t)$, le signe $\equiv$ indiquant ainsi que les deux membres sont égaux pour tout t de l'intervalle $[a, b]$. Considérons les deux opérations suivantes:

- $\mathbf{f} + \mathbf{g}$, définie par la formule $(\mathbf{f} + \mathbf{g})(t) \equiv \mathbf{f}(t) + \mathbf{g}(t)$,
- $\alpha\mathbf{f}$, définie par la formule $(\alpha\mathbf{f})(t) \equiv \alpha\mathbf{f}(t)$.

Ces deux opérations satisfont aux conditions (a)–(h) de la définition 1.2.2 et munissent $C_{[a,b]}$ d'une structure d'espace vectoriel. Le vecteur nul de cet espace est la fonction nulle et l'opposé de $\mathbf{f}$ est la fonction $-\mathbf{f}$ définie par $(-\mathbf{f})(t) \equiv -\mathbf{f}(t)$.

Il est intéressant de constater que $C_{[a,b]}$, en tant qu'espace vectoriel, est une généralisation naturelle de $\mathbb{R}^n$ au cas continu. On peut en effet concevoir tout vecteur (a_i) de $\mathbb{R}^n$ sous la forme d'une fonction réelle définie dans l'ensemble $\{1, 2, ..., n\}$: la valeur de cette fonction au point i est tout simplement a_i.

1.3.5 Autres espaces vectoriels fonctionnels

Voici quelques autres exemples d'espaces vectoriels fonctionnels. Les opérations d'addition et de multiplication par un scalaire sont définies comme dans 1.3.4.

(1) L'espace vectoriel $C^k_{(a,b)}$ formé des fonctions réelles k fois continûment dérivables, définies dans l'intervalle ouvert (a, b).

(2) L'espace vectoriel des fonctions réelles définies dans un intervalle.

(3) L'espace vectoriel des polynômes à une indéterminée.

(4) L'espace vectoriel des polynômes à une indéterminée de degré inférieur ou égal à n.

1.4 COMBINAISONS LINÉAIRES, SOUS-ESPACES VECTORIELS, FAMILLES GÉNÉRATRICES

1.4.1 Avertissement

Dorénavant, sauf mention explicite du contraire, les vecteurs seront les éléments d'un espace vectoriel donné E.

1.4.2 Combinaisons linéaires

On appelle *combinaison linéaire* des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ tout vecteur de la forme $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k$, où $\alpha_1, \alpha_2, ..., \alpha_k$ sont des nombres appelés *coefficients* de la combinaison linéaire.

1.4.3 Exemples

(1) Le vecteur nul est combinaison linéaire de $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$. Pour voir cela, il suffit de prendre $\alpha_1 = \alpha_2 = ... = \alpha_k = 0$. Dans ce cas, la combinaison linéaire est appelée *combinaison linéaire triviale*.

(2) Les combinaisons linéaires d'un vecteur $\mathbf{x}$ sont appelées *multiples* de $\mathbf{x}$. Un multiple de $\mathbf{x}$ est donc un vecteur de la forme $\alpha\mathbf{x}$. On notera que le vecteur nul est multiple de tout vecteur.

(3) *Combinaisons convexes.* On appelle *combinaison convexe* toute combinaison linéaire dont les coefficients sont non négatifs et de somme égale à 1. L'ensemble des combinaisons convexes de deux points P et Q de $\mathscr{E}_O$ est le segment de droite joignant P et Q. Pour s'en rendre compte, il suffit d'écrire

$$\alpha P + (1 - \alpha)Q = Q + \alpha(P - Q),$$

de faire varier α de 0 à 1 et de constater que tous les points du segment sont ainsi obtenus (fig. 1.9).

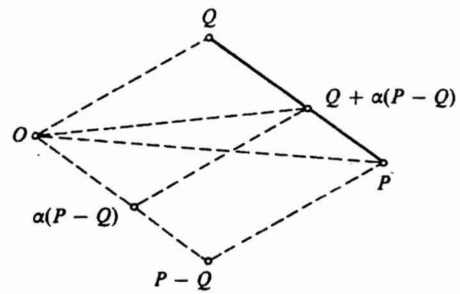

Fig. 1.9

Nous laissons au lecteur le soin de vérifier que l'ensemble des combinaisons convexes de trois points est le triangle, éventuellement dégénéré, dont les sommets sont ces trois points.

D'une manière générale, on peut montrer que l'ensemble des combinaisons convexes de $k > 3$ points est le plus petit polyèdre convexe (polygone convexe si $\mathscr{G}_O$ désigne le plan), éventuellement dégénéré, comprenant ces points.

(4) On dit que le vecteur-colonne (x_i^0) de $\mathbb{R}^3$ est solution du système d'équations linéaires

$$\begin{aligned} 3x_1 - 2x_2 + 4x_3 &= 0 \\ -x_1 + x_2 - 5x_3 &= 0 \end{aligned} \tag{1.2}$$

si ses termes x_1^0, x_2^0, x_3^0, substitués aux inconnues x_1, x_2, x_3, vérifient les deux équations.

Toute combinaison linéaire $\alpha_1(x_i^1) + \alpha_2(x_i^2)$ de solutions (x_i^1) et (x_i^2) du système (1.2) est encore une solution de ce système.

(5) Soit $\mathbf{f}_1$ et $\mathbf{f}_2$ les deux fonctions définies par

$$\mathbf{f}_1(t) \equiv \cos t \quad \text{et} \quad \mathbf{f}_2(t) \equiv \sin t. \tag{1.3}$$

Toute combinaison linéaire de $\mathbf{f}_1$ et $\mathbf{f}_2$ est solution de l'équation différentielle

$$\ddot{\mathbf{f}} + \mathbf{f} = \mathbf{0},$$

où $\ddot{\mathbf{f}}$ désigne la deuxième dérivée de $\mathbf{f}$.

1.4.4 Combinaisons linéaires itérées

Si le vecteur $\mathbf{x}$ est combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ et chacun de ces vecteurs est combinaison linéaire des vecteurs $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l$, alors $\mathbf{x}$ est combinaison linéaire de $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l$.

1.4.5 Sous-espaces vectoriels

On appelle *sous-espace vectoriel* de E tout sous-ensemble de E qui est lui-même un espace vectoriel pour les opérations d'addition et de multiplication par un scalaire définies dans E.

Un sous-espace vectoriel, en tant qu'espace vectoriel, ne peut être vide, puisqu'il comprend au moins un vecteur, à savoir son vecteur nul, celui-ci étant d'ailleurs forcément le vecteur nul de E, en vertu de la règle (5) de 1.2.3. En outre, en même temps que les vecteurs $\mathbf{x}$ et $\mathbf{y}$, il comprend toutes leurs combinaisons linéaires $\alpha\mathbf{x} + \beta\mathbf{y}$. Inversement, on voit aussitôt que tout sous-ensemble jouissant de ces propriétés est un sous-espace vectoriel. Nous avons ainsi établi la proposition suivante:

1.4.6 Proposition. Caractérisation des sous-espaces vectoriels

Un sous-ensemble S de E est un sous-espace vectoriel de E si et seulement si S est non vide et $\alpha\mathbf{x} + \beta\mathbf{y}$ appartient à S pour tout couple $(\mathbf{x}, \mathbf{y})$ *de vecteurs de S et tout couple (α, β) de nombres.*

La proposition suivante en découle aisément:

1.4.7 Proposition

Soit $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ une famille de vecteurs. L'ensemble des combinaisons linéaires de $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ est un sous-espace vectoriel S de E, plus précisément le plus petit sous-espace vectoriel de E (au sens de l'inclusion) comprenant $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$.

1.4.8 Générateurs, familles génératrices

Les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ de la proposition 1.4.7 sont appelés *générateurs* de S et la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ *famille génératrice* de S. On dit aussi que ces vecteurs ou cette famille *engendrent* S.

1.4.9 Somme et intersection de sous-espaces vectoriels

Soit S et T des sous-espaces vectoriels de E. On appelle *somme* de S et T, et on note $S + T$, l'ensemble des vecteurs de la forme $\mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ et un vecteur de S et $\mathbf{t}$ un vecteur de T. A l'aide de la proposition 1.4.7, le lecteur constatera facilement que la somme $S + T$ et l'intersection $S \cap T$ sont des sous-espaces vectoriels de E. Au contraire, la réunion $S \cup T$ n'est pas un sous-espace vectoriel de E, à moins que S ne soit contenu dans T, ou réciproquement. En fait, le plus petit sous-espace vectoriel de E contenant $S \cup T$ est la somme $S + T$ (cf. exercice 1.12.14).

Les résultats de ce paragraphe s'étendent, de manière évidente, au cas d'une famille quelconque $(S_1, S_2, ..., S_k)$ de sous-espaces vectoriels de E: la somme $S_1 + S_2 + ... + S_k$, c'est-à-dire l'ensemble des vecteurs de la forme $\mathbf{s}_1 + \mathbf{s}_2 + ... + \mathbf{s}_k$, où $\mathbf{s}_i$ est un vecteur de S_i pour $i = 1, 2, ..., k$, et l'intersection $S_1 \cap S_2 \cap ... \cap S_k$ sont des sous-espaces vectoriels de E.

1.4.10 Exemples

(1) $\{\mathbf{0}\}$ et E sont des sous-espaces vectoriels de E.

(2) Le sous-espace vectoriel engendré par un vecteur non nul est formé de tous les multiples de ce vecteur. On appelle un tel sous-espace *droite vectorielle*. Un sous-espace vectoriel engendré par deux vecteurs non multiples l'un de l'autre est appelé *plan vectoriel*. Dans $\mathscr{G}_O$ une droite et un plan vectoriels sont effectivement une droite et un plan passant par l'origine O.

(3) $\mathbb{R}^4$ est engendré par les vecteurs-colonnes

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ \begin{pmatrix}0\\0\\1\\0\end{pmatrix},\ \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

En effet,

$$\begin{pmatrix}a_1\\a_2\\a_3\\a_4\end{pmatrix} = a_1\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + a_2\begin{pmatrix}0\\1\\0\\0\end{pmatrix} + a_3\begin{pmatrix}0\\0\\1\\0\end{pmatrix} + a_4\begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{1.4}$$

(4) Les fonctions $\mathbf{f}_1$ et $\mathbf{f}_2$ définies dans (1.3) engendrent le sous-espace vectoriel de $C^2_{(a,b)}$ formé des solutions de l'équation différentielle $\ddot{\mathbf{f}} + \mathbf{f} = \mathbf{0}$. Ce résultat d'analyse est énoncé sans démonstration.

(5) Etant donné cinq nombres $t_0, \ldots, t_4$ arbitrairement choisis, définissons les cinq polynômes du quatrième degré $\mathbf{p}_0, \ldots, \mathbf{p}_4$ par la formule

$$\mathbf{p}_i(t) \equiv \frac{(t-t_0)\ldots(t\,\hat{-}\,t_i)\ldots(t-t_4)}{(t_i-t_0)\ldots(t_i\,\hat{-}\,t_i)\ldots(t_i-t_4)}, \tag{1.5}$$

où l'accent circonflexe indique l'absence des facteurs d'indice i. Il est évident que

$$\mathbf{p}_i(t_i) = 1 \quad \text{et} \quad \mathbf{p}_i(t_j) = 0 \text{ si } i \neq j. \tag{1.6}$$

Nous allons établir que ces cinq polynômes engendrent l'espace vectoriel de tous les polynômes de degré inférieur ou égal à 4, en démontrant que si $\mathbf{p}$ est un tel polynône, alors

$$\mathbf{p}(t) \equiv \mathbf{p}(t_0)\,\mathbf{p}_0(t) + \ldots + \mathbf{p}(t_4)\,\mathbf{p}_4(t) \quad \textit{(formule de Lagrange)}. \tag{1.7}$$

A cet effet, désignons par $\tilde{\mathbf{p}}$ le polynôme défini par le second membre de (1.7). D'après (1.6),

$$\tilde{\mathbf{p}}(t_0) = \mathbf{p}(t_0)\cdot 1 + 0 + 0 + 0 + 0 = \mathbf{p}(t_0)$$
$$\ldots\ldots$$
$$\tilde{\mathbf{p}}(t_4) = 0 + 0 + 0 + 0 + \mathbf{p}(t_4)\cdot 1 = \mathbf{p}(t_4),$$

ce qui montre que $\mathbf{p}$ et $\tilde{\mathbf{p}}$ prennent les mêmes valeurs en $t = t_0, \ldots, t = t_4$, donc que le polynôme $\tilde{\mathbf{p}} - \mathbf{p}$ a au moins cinq zéros. Puisqu'il est de degré inférieur ou égal à 4, nous en concluons qu'il est nul, donc que $\tilde{\mathbf{p}} = \mathbf{p}$. La formule (1.7) est ainsi démontrée.

(6) L'ensemble des polynômes de degré inférieur ou égal à n est un sous-espace vectoriel de l'espace de tous les polynômes. Il est engendré par les monômes

$$\mathbf{p}_0\colon \mathbf{p}_0(t) \equiv 1\,,\quad \mathbf{p}_1\colon \mathbf{p}_1(t) \equiv t\,,\quad \ldots,\quad \mathbf{p}_n\colon \mathbf{p}_n(t) \equiv t^n. \tag{1.8}$$

(7) Si $\mathbf{x}$ n'est pas multiple de $\mathbf{y}$, $S = \{\mathbf{z}\colon \mathbf{z} = \mathbf{x} + \alpha\mathbf{y}, \alpha \in \mathbb{R}\}$ n'est pas un sous-espace vectoriel de E, ne serait-ce que par le fait que $\mathbf{0}$ n'est pas élément de S. En particulier, une droite ne passant pas par l'origine n'est pas un sous-espace vectoriel de $\mathscr{G}_O$. (Pour plus de détails, voir les sections 1.8 et 1.9.)

1.5 DÉPENDANCE ET INDÉPENDANCE LINÉAIRES

1.5.1 Caractérisation de l'absence de parallélisme à un même plan

Si $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ sont trois vecteurs de V^3 dont les représentants ne sont pas parallèles à un même plan (par convention, une flèche d'intensité nulle est parallèle à tout plan), alors tout vecteur $\mathbf{x}$ de V^3 s'écrit de manière unique sous la forme

$$\mathbf{x} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3,$$

où α_1, α_2, α_3 sont des nombres (fig. 1.10).

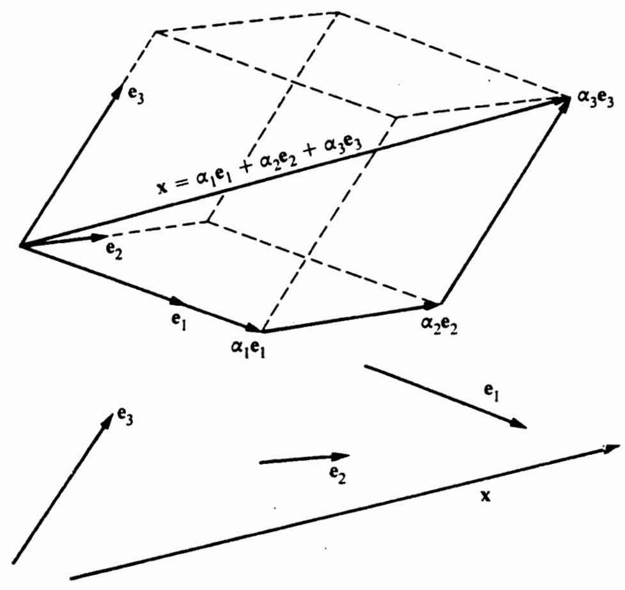

Fig. 1.10

En particulier, la seule possibilité d'obtenir le vecteur nul comme combinaison linéaire de $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ est d'attribuer la valeur 0 à α_1, α_2, α_3.

Réciproquement, si pour trois vecteurs $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ de V^3 la relation $\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3 = \mathbf{0}$ implique $\alpha_1 = \alpha_2 = \alpha_3 = 0$, aucun de ces vecteurs ne peut être

combinaison linéaire des deux autres, autrement dit, leurs représentants ne sont pas parallèles à un même plan.

Sur la base de ces observations, nous allons étendre la notion d'absence de parallélisme à un même plan au cas d'un nombre quelconque de vecteurs d'un espace vectoriel E.

1.5.2 Indépendance et dépendance linéaires, familles libres et liées

On dit que les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont *linéairement indépendants* si la relation $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$ implique $\alpha_1 = \alpha_2 = ... = \alpha_k = 0$, autrement dit, si la combinaison linéaire triviale est la seule combinaison linéaire de $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ qui soit nulle. Dans le cas contraire, on dit que les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont *linéairement dépendants*.

Si l'attention est fixée sur la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ plutôt que sur les termes dont elle est constituée, on dit que celle-ci est *libre* ou *liée* suivant que les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement indépendants ou dépendants.

1.5.3 Formulation de la dépendance linéaire

Selon la définition 1.5.2, les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement dépendants s'il existe des nombres non tous nuls $\alpha_1, \alpha_2, ..., \alpha_k$ tels que $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$.

1.5.4 Exemples

Voici quelques cas où la définition 1.5.2 s'applique directement.

(1) Une famille réduite à un seul terme $\mathbf{x}$ est libre ou liée suivant que $\mathbf{x}$ est non nul ou nul.

(2) Pour qu'un couple $(\mathbf{x}, \mathbf{y})$ soit lié, il faut et il suffit que l'un des vecteurs $\mathbf{x}$ ou $\mathbf{y}$ soit multiple de l'autre. On remarquera que le couple $(\mathbf{x}, \mathbf{0})$ est lié, mais que $\mathbf{x}$ n'est pas multiple de $\mathbf{0}$ si $\mathbf{x} \neq \mathbf{0}$.

(3) Si un des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ est nul, ces vecteurs sont linéairement dépendants, car, en supposant $\mathbf{x}_i = \mathbf{0}$, nous voyons que la combinaison linéaire $0\mathbf{x}_1 + ... + 1\mathbf{x}_i + ... + 0\mathbf{x}_k$ est nulle sans être triviale.

(4) Si $\mathbf{x}_i = \mathbf{x}_j$ pour un couple d'indices i et j différents, alors les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement dépendants, car, étant admis par exemple que $i < j$, la combinaison linéaire $0\mathbf{x}_1 + ... + 1\mathbf{x}_i + ... + (-1)\mathbf{x}_j + ... + 0\mathbf{x}_k$ est nulle sans être triviale. Cet exemple et le précédent se généralisent de la façon suivante:

(5) Si les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement dépendants, alors les vecteurs $\mathbf{x}_1, ..., \mathbf{x}_k, \mathbf{x}_{k+1}, ..., \mathbf{x}_l$ le sont aussi, quels que soient $\mathbf{x}_{k+1}, ..., \mathbf{x}_l$ $(l > k)$. En d'autres termes, toute famille finie de vecteurs admettant une sous-famille liée est elle-même liée.

(6) Si les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement indépendants, les vecteurs $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, ..., \mathbf{x}_{i_l}$ $(1 \leqslant i_1 < i_2 < ... < i_l \leqslant k)$ le sont également. En d'autres termes, toute sous-famille d'une famille libre est elle-même libre.

La proposition suivante généralise l'exemple (2):

1.5.5 Proposition. Caractérisation de la dépendance linéaire

Pour qu'une famille de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ $(k > 1)$ soit liée, il faut et il suffit qu'un vecteur $\mathbf{x}_i$ soit combinaison linéaire des vecteurs $\mathbf{x}_j$ avec $j \neq i$.

DÉMONSTRATION

Si la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est liée, il existe des nombres non tous nuls $\alpha_1, \alpha_2, ..., \alpha_k$ tels que $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$. En supposant α_i non nul et en résolvant par rapport à $\mathbf{x}_i$, nous obtenons

$$\mathbf{x}_i = \frac{1}{\alpha_i}(-\alpha_1\mathbf{x}_1 - ... \widehat{\alpha_i\mathbf{x}_i} - ... - \alpha_k\mathbf{x}_k),$$

où l'accent circonflexe indique l'absence du terme d'indice i. Cela montre que $\mathbf{x}_i$ est combinaison linéaire des $\mathbf{x}_j$ avec $j \neq i$.

Inversement, si $\mathbf{x}_i = \alpha_1\mathbf{x}_1 + ... + \widehat{\alpha_i\mathbf{x}_i} + ... + \alpha_k\mathbf{x}_k$, alors $\alpha_1\mathbf{x}_1 + ... + (-1)\mathbf{x}_i + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$, ce qui montre que la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est liée, car au moins un coefficient de la combinaison linéaire est non nul.

1.5.6 Proposition. Critère d'indépendance linéaire

Pour qu'une famille de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ soit libre, il faut et il suffit qu'aucun vecteur $\mathbf{x}$ ne puisse s'écrire de deux manières sous la forme d'une combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$.

DÉMONSTRATION

Supposons que la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ soit libre. S'il existait un vecteur $\mathbf{x}$ tel que

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \alpha'_1\mathbf{x}_1 + \alpha'_2\mathbf{x}_2 + ... + \alpha'_k\mathbf{x}_k,$$

avec $\alpha_i \neq \alpha'_i$ pour au moins un indice i, la combinaison linéaire

$$(\alpha_1 - \alpha'_1)\mathbf{x}_1 + (\alpha_2 - \alpha'_2)\mathbf{x}_2 + ... + (\alpha_k - \alpha'_k)\mathbf{x}_k$$

serait nulle sans être triviale, ce qui contredirait l'hypothèse.

Réciproquement, supposons qu'aucun vecteur $\mathbf{x}$ ne puisse être écrit de deux manières sous la forme d'une combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, .., \mathbf{x}_k$. Alors, en particulier, le vecteur nul ne pourra l'être, autrement dit, la combinaison linéaire triviale est la seule combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ qui s'annule, donc la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est libre.

1.6 BASES D'UN ESPACE VECTORIEL

1.6.1 Introduction

Les familles génératrices libres jouent un rôle important en algèbre linéaire. Elles seront étudiées dans la présente section et la suivante.

Comme précédemment, E désignera un espace vectoriel.

1.6.2 Bases d'un espace vectoriel

On dit qu'une famille finie de vecteurs est une *base* de E si elle est libre et engendre E.

D'après cette définition, toute famille libre $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est une base du sous-espace vectoriel qu'elle engendre.

1.6.3 Proposition

Pour qu'une famille de vecteurs $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ *soit une base de E, il faut et il suffit que tout vecteur* $\mathbf{x}$ *de E s'exprime de manière unique sous la forme d'une combinaison linéaire des vecteurs* $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n. \tag{1.9}$$

L'expression (1.9) est appelée *décomposition* de $\mathbf{x}$ suivant la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

DÉMONSTRATION

Par définition d'une base, les vecteurs $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ engendrent E, donc tout vecteur $\mathbf{x}$ s'écrit sous la forme (1.9). D'autre part, puisque $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ sont linéairement indépendants, la décomposition (1.9) est unique, d'après la proposition 1.5.6.

Réciproquement, si tout vecteur $\mathbf{x}$ s'écrit de manière unique sous la forme (1.9), les vecteurs $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ engendrent E et, en outre, ils sont linéairement indépendants, encore d'après la proposition 1.5.6, donc $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est une base de E.

1.6.4 Composantes d'un vecteur

Les coefficients $x_1, x_2, ..., x_n$ de la décomposition d'un vecteur $\mathbf{x}$ suivant une base sont appelés *composantes* de $\mathbf{x}$ dans cette base.

En présence d'une base, tout vecteur est donc entièrement déterminé par ses composantes.

1.6.5 Proposition

Si $x_1, x_2, ..., x_n$ *sont les composantes de* $\mathbf{x}$ *et* $y_1, y_2, ..., y_n$ *celles de* $\mathbf{y}$, *alors* $x_1 + y_1, x_2 + y_2, ..., x_n + y_n$ *sont les composantes de* $\mathbf{x} + \mathbf{y}$ *et* $\alpha x_1, \alpha x_2, ..., \alpha x_n$ *celles de* $\alpha\mathbf{x}$.

En d'autres termes, additionner deux vecteurs revient à additionner leurs composantes et multiplier un vecteur par α revient à multiplier ses composantes par α. La base est donc un outil de calcul important, car elle permet d'effectuer les opérations sur les vecteurs au moyen d'opérations sur les nombres.

1.6.6 Notation

De toute évidence, les composantes d'un vecteur dépendent du choix de la base, mais une fois que ce choix est fait, il n'y aura aucun danger d'ambiguïté lorsqu'on écrira $\mathbf{x}(x_1, x_2, ..., x_n)$ pour exprimer que $x_1, x_2, ..., x_n$ sont les composantes de $\mathbf{x}$. (Une écriture mieux appropriée au cas où des changements de base interviennent sera introduite dans 6.5.6.)

1.6.7 Exemples

(1) La donnée d'une base dans $\mathscr{G}_O$ équivaut à celle d'un système d'axes de référence d'origine O dans $\mathscr{G}$. Les composantes d'un point de $\mathscr{G}_O$ sont, dans ce cas, les coordonnées de ce point par rapport au système d'axes.

(2) Deux (trois) vecteurs linéairement indépendants de V^2 (V^3) forment une base.

(3) Les vecteurs-colonnes de $\mathbb{R}^n$

$$\begin{bmatrix}1\\0\\\cdot\\\cdot\\\cdot\\0\end{bmatrix}, \begin{bmatrix}0\\1\\\cdot\\\cdot\\\cdot\\0\end{bmatrix}, ..., \begin{bmatrix}0\\0\\\cdot\\\cdot\\\cdot\\1\end{bmatrix} \tag{1.10}$$

forment une base que l'on appelle *base canonique* de $\mathbb{R}^n$. Les composantes du vecteur-colonne (a_i) dans cette base sont $a_1, a_2, ..., a_n$.

(4) Les polynômes $\mathbf{p}_0, ..., \mathbf{p}_4$ définis dans (1.5) forment une base de l'espace vectoriel des polynômes de degré inférieur ou égal à 4. En effet, ils engendrent cet espace et, de plus, ils sont linéairement indépendants, car la relation $\alpha_0\mathbf{p}_0 + ... + \alpha_4\mathbf{p}_4 = \mathbf{0}$ implique $\alpha_0\mathbf{p}_0(t_i) + ... + \alpha_4\mathbf{p}_4(t_i) = 0$ et donc, par (1.6), $\alpha_i = 0$ pour $i = 0, ..., 4$. D'après (1.7), les composantes dans cette base d'un polynôme $\mathbf{p}$ de degré inférieur ou égal à 4 sont $\mathbf{p}(t_0), ..., \mathbf{p}(t_4)$.

(5) Les monômes $\mathbf{p}_0, ..., \mathbf{p}_n$ définis dans (1.8) forment une base de l'espace vectoriel des polynômes de degré inférieur ou égal à n. En effet, ils engendrent cet

espace et, en outre, ils sont linéairement indépendants, car la relation $\alpha_0\mathbf{p}_0 + \ldots + \alpha_n\mathbf{p}_n = \mathbf{0}$ s'écrit aussi sous la forme $\alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n \equiv 0$ et implique donc $\alpha_0 = \alpha_1 = \ldots = \alpha_n = 0$ (cf. exercice 1.12.5). Les composantes d'un polynôme de degré inférieur ou égal à n dans cette base sont les coefficients de ce polynôme.

1.7 DIMENSION D'UN ESPACE VECTORIEL

1.7.1 Introduction

Au long de ce livre, nous nous occuperons principalement d'espaces vectoriels admettant une base. Un des objectifs de cette section est de caractériser ces espaces.

Dans cette section, E désignera encore un espace vectoriel.

1.7.2 Dimension finie et infinie

On dit que E est de *dimension finie* s'il est engendré par une famille finie de vecteurs. Dans le cas contraire, on dit que E est de *dimension infinie*.

1.7.3 Théorème. Prolongement d'une famille libre

Soit $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *une famille libre et* $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ *une famille génératrice de E. Si* $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *n'est pas une base de E, on peut extraire une sous-famille* $(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_j})$ *de* $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ *de telle manière que la famille* $(\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_j})$ *soit une base de E.*

DÉMONSTRATION

Au moins un des vecteurs $\mathbf{v}_i$ n'est pas combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, sinon $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ engendrerait E, en raison de 1.4.4, et serait donc une base de E. Notons ce vecteur $\mathbf{v}_{i_1}$. La famille $(\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{v}_{i_1})$ est alors libre. En effet, la relation $\alpha_1\mathbf{x}_1 + \ldots + \alpha_k\mathbf{x}_k + \beta_1\mathbf{v}_{i_1} = \mathbf{0}$ implique d'abord $\beta_1 = 0$, autrement $\mathbf{v}_{i_1}$ serait combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, et ensuite $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, puisque les vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ sont linéairement indépendants. Si la famille $(\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{v}_{i_1})$ engendre E, elle est une base de E et le théorème est alors démontré. Dans le cas contraire, le même raisonnement nous assure de l'existence d'un vecteur $\mathbf{v}_{i_2}$ tel que $(\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{v}_{i_1}, \mathbf{v}_{i_2})$ est une famille libre. Si cette famille n'engendre pas E, le procédé d'extraction de vecteurs $\mathbf{v}_i$ de $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ se poursuit. Lorsqu'il s'arrête, nous aurons obtenu un prolongement de $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ en une famille libre engendrant E, c'est-à-dire en une base de E.

1.7.4 Corollaire. Existence et extraction d'une base

Tout espace vectoriel de dimension finie et non réduit au vecteur nul admet une base. En fait, de toute famille génératrice d'un tel espace on peut extraire une base.

DÉMONSTRATION

Soit $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ une famille génératrice de E. Si E n'est pas réduit à $\{\mathbf{0}\}$, au moins un vecteur $\mathbf{v}_i$ n'est pas nul. Désignons ce vecteur par $\mathbf{x}$. Le corollaire résulte alors du théorème 1.7.3 appliqué aux familles $(\mathbf{x})$ et $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$.

1.7.5 Théorème

Si $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *est une base de E, toute famille de vecteurs* $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *dont le nombre de termes k est supérieur à n est liée.*

DÉMONSTRATION

Nous raisonnons par l'absurde en supposant que la famille $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ soit libre. Considérons la décomposition de $\mathbf{x}_1$ suivant la base $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$:

$$\mathbf{x}_1 = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \ldots + \alpha_n\mathbf{e}_n.$$

Comme $\mathbf{x}_1$ n'est pas nul, au moins un des coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ n'est pas nul. Quitte à énumérer autrement les termes de la base, nous pouvons admettre que ce coefficient est α_1. Dans ce cas,

$$\mathbf{e}_1 = \frac{1}{\alpha_1}\mathbf{x}_1 - \frac{\alpha_2}{\alpha_1}\mathbf{e}_2 - \ldots - \frac{\alpha_n}{\alpha_1}\mathbf{e}_n$$

et donc $(\mathbf{x}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ est une famille génératrice de E, puisque le sous-espace vectoriel qu'elle engendre comprend les vecteurs $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Exprimons alors $\mathbf{x}_2$ sous la forme d'une combinaison linéaire de $\mathbf{x}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$:

$$\mathbf{x}_2 = \beta_1\mathbf{x}_1 + \beta_2\mathbf{e}_2 + \ldots + \beta_n\mathbf{e}_n.$$

Les coefficients $\beta_2, \ldots, \beta_n$ ne sont pas tous nuls, sinon $\mathbf{x}_1$ et $\mathbf{x}_2$ seraient linéairement dépendants. Sans restreindre la généralité, nous pouvons admettre que β_2 n'est pas nul, ce qui nous permet d'écrire

$$\mathbf{e}_2 = -\frac{\beta_1}{\beta_2}\mathbf{x}_1 + \frac{1}{\beta_2}\mathbf{x}_2 - \frac{\beta_3}{\beta_2}\mathbf{e}_3 - \ldots - \frac{\beta_n}{\beta_2}\mathbf{e}_n$$

et donc de conclure que $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n)$ est une famille génératrice de E, puisque le sous-espace vectoriel qu'elle engendre comprend les vecteurs $\mathbf{x}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. La suite de la démonstration poursuit ce procédé d'échange: $\mathbf{e}_3, \ldots, \mathbf{e}_n$ sont remplacés successivement par $\mathbf{x}_3, \ldots, \mathbf{x}_n$. Le résultat montre que la famille $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ engendre E. Mais alors $\mathbf{x}_k$ est combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, ce qui contredit l'hypothèse, en raison de la proposition 1.5.5.

1.7.6 Corollaire

Si $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ *et* $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_k)$ *sont deux bases de E, alors* $n = k$.

DÉMONSTRATION

Du théorème 1.7.5 nous déduisons que $k \leqslant n$, ainsi que $n \leqslant k$, par un échange du rôle des deux bases. Il s'ensuit que $n = k$.

1.7.7 Dimension d'un espace vectoriel

Au moyen des corollaires 1.7.4 et 1.7.6, il est maintenant possible d'attribuer une dimension à tout espace vectoriel de dimension finie. Soit E un tel espace. On appelle *dimension* de E le nombre de termes d'une quelconque de ses bases. Si E se réduit au seul vecteur nul, on dit que sa dimension est nulle. La dimension de E sera notée $\dim E$.

1.7.8 Exemples

La dimension de V^2 est 2, celle de V^3 est 3 et celle de $\mathbb{R}^n$ est n. D'après l'exemple (5) de 1.6.7, la dimension de l'espace vectoriel des polynômes de degré inférieur ou égal à n est $n + 1$. L'espace vectoriel de tous les polynômes et, par conséquent, les espaces vectoriels $C_{[a, b]}$ et $C^k_{(a, b)}$ sont de dimension infinie. En effet, ces espaces admettent des familles libres arbitrairement grandes, par exemple $(\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_n)$, où les $\mathbf{p}_i$ sont les monômes définis dans (1.8), n étant pris arbitrairement grand; par le théorème 1.7.5, ces espaces n'admettent aucune base, donc leur dimension est infinie, d'après le corollaire 1.7.4.

1.7.9 Proposition. Caractérisations d'une base

Supposons que E soit de dimension finie non nulle n.
(a) *Toute famille libre à n termes est une base de E.*
(b) *Toute famille génératrice à n termes est une base de E.*

DÉMONSTRATION

Assertion (a). Une famille libre à n termes qui ne serait pas une base se prolongerait en une base, d'après le théorème 1.7.3, et la dimension de E serait alors supérieure à n.

Assertion (b). D'une famille génératrice à n termes qui ne serait pas une base on pourrait extraire une base, d'après le corollaire 1.7.4, et la dimension de E serait alors inférieure à n.

1.7.10 Isomorphie de E et $\mathbb{R}^n$

Tout espace vectoriel E de dimension finie non nulle n peut être mis en correspondance biunivoque avec $\mathbb{R}^n$. Il suffit de choisir une base de E et de faire correspondre à tout vecteur $\mathbf{x}$ de E le vecteur-colonne dont les termes sont les composantes de $\mathbf{x}$ dans la base choisie:

$$\begin{array}{ccc} E & & \mathbb{R}^n \\ \mathbf{x}(x_1, x_2, ..., x_n) & \longleftrightarrow & \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \end{array} . \tag{1.11}$$

D'après la proposition 1.6.5, cette correspondance conserve les opérations d'addition vectorielle et de multiplication par un scalaire; en d'autres termes, elle permet, comme nous l'avons déjà fait remarquer, d'effectuer les opérations sur les vecteurs par des opérations sur les nombres. On dit que E et $\mathbb{R}^n$ sont *isomorphes*, ou que la correspondance est un *isomorphisme* (cf. 6.3.7). Evidemment, cet isomorphisme dépend de la base de E choisie.

Il importe de noter qu'un espace vectoriel E de dimension n n'admet, en général, aucune base privilégiée, contrairement à $\mathbb{R}^n$; par conséquent, sauf si le choix d'une base de E a été fait, il faudra éviter de considérer E et $\mathbb{R}^n$ comme identiques.

1.8 RETOUR AUX SOUS-ESPACES VECTORIELS, SOMMES DIRECTES

1.8.1 Introduction

Dans cette section, nous reprenons l'étude des sous-espaces vectoriels d'un espace vectoriel donné E et introduisons les notions de rang d'une famille finie de vecteurs, ainsi que celle de somme directe de sous-espaces vectoriels.

1.8.2 Rang d'une famille finie de vecteurs

On appelle *rang* d'une famille finie de vecteurs la dimension du sous-espace vectoriel de E qu'elle engendre.

1.8.3 Proposition

Le rang d'une famille de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ *est inférieur ou égal à k. Il est égal à k si et seulement si cette famille est libre.*

DÉMONSTRATION

Ecartons le cas trivial où le rang de la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est nul. D'après le corollaire 1.7.4, on peut alors extraire de cette famille une base du sous-espace vectoriel qu'elle engendre. Le rang est donc inférieur à k ou égal à k suivant que $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est une famille liée ou libre.

1.8.4 Proposition. Comparaison de deux rangs

Pour que le rang d'une famille de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ *soit égal au rang de la famille augmentée* $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k, \mathbf{y})$, *il faut et il suffit que le vecteur* $\mathbf{y}$ *soit combinaison linéaire des vecteurs* $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$.

DÉMONSTRATION

Notons S et T les sous-espaces vectoriels engendrés respectivement par $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ et $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k, \mathbf{y})$. Si $\mathbf{y}$ est combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$, alors $S = T$, donc les deux rangs sont égaux. Réciproquement, supposons que les deux rangs soient égaux et montrons que $S = T$, ce qui nous permettra de conclure que $\mathbf{y}$ est combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$. Si les deux rangs sont nuls, il est clair que $S = T = \{\mathbf{0}\}$. Sinon, une base de S est également une base de T, puisque S est inclus dans T, ce qui entraîne que $S = T$.

1.8.5 Proposition. Sous-espaces vectoriels d'un espace vectoriel de dimension finie

Si E est de dimension finie et S est un sous-espace vectoriel de E, alors S est de dimension finie et $\dim S \leqslant \dim E$. *En outre,* $\dim S = \dim E$ *si et seulement si* $S = E$.

DÉMONSTRATION

Désignons la dimension de E par n et supposons que celle de S soit infinie ou finie et supérieure à n. Par récurrence, nous allons démontrer l'existence d'une famille libre $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{n+1})$ de vecteurs de S, ce qui contredit le théorème 1.7.5, vu la définition 1.7.7. Comme S n'est pas de dimension nulle, il comprend au moins un vecteur non nul $\mathbf{x}_1$, donc $(\mathbf{x}_1)$ est une famille libre. Supposons que $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est une famille libre de vecteurs de S. Si k est inférieur ou égal à n, au moins un des vecteurs de S n'est pas combinaison linéaire de $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$, sinon S serait de dimension k, contrairement à l'hypothèse. Désignons ce vecteur par $\mathbf{x}_{k+1}$. En vertu de la proposition précédente, la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{k+1})$ de vecteurs de S est alors libre.

Pour démontrer la seconde assertion, il suffit de poser $T = E$ dans la dernière partie de la démonstration précédente.

1.8.6 Hyperplans vectoriels

Si E est de dimension finie non nulle n, un sous-espace vectoriel de E de dimension $n - 1$ est appelé *hyperplan vectoriel*. Par exemple, un hyperplan vectoriel se réduit au vecteur nul si $n = 1$, est une droite vectorielle si $n = 2$, un plan vectoriel si $n = 3$.

1.8.7 Sommes directes

On dit que la somme $S + T$ de deux sous-espaces vectoriels S et T de E est *directe* si $S \cap T = \{\mathbf{0}\}$. Dans ce cas, on la note $S \oplus T$.

Par exemple, si E est de dimension 2 et $(\mathbf{e}_1, \mathbf{e}_2)$ est une base de E, alors $E = D_1 \oplus D_2$, où D_1 et D_2 sont les droites vectorielles engendrées respectivement par $\mathbf{e}_1$ et $\mathbf{e}_2$.

• Tout vecteur d'une somme directe $S \oplus T$ s'écrit d'une seule manière sous la forme $\mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T. En effet, si

$$\mathbf{s} + \mathbf{t} = \mathbf{u} + \mathbf{v},$$

où $\mathbf{s}$, $\mathbf{u}$ sont des vecteurs de S et $\mathbf{t}$, $\mathbf{v}$ des vecteurs de T, alors

$$\mathbf{s} - \mathbf{u} = \mathbf{v} - \mathbf{t}$$

est un vecteur de $S \cap T = \{\mathbf{0}\}$, donc $\mathbf{s} - \mathbf{u} = \mathbf{v} - \mathbf{t} = \mathbf{0}$, ce qui entraîne $\mathbf{s} = \mathbf{u}$ et $\mathbf{t} = \mathbf{v}$.

Inversement, si tout vecteur d'une somme $S + T$ s'écrit d'une seule manière sous la forme $\mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T, alors $S \cap T = \{\mathbf{0}\}$ et donc $S + T$ est une somme directe. En effet, un vecteur non nul $\mathbf{u}$ de $S \cap T$ permettrait au vecteur nul d'avoir deux décompositions, à savoir $\mathbf{0} = \mathbf{0} + \mathbf{0}$ et $\mathbf{0} = \mathbf{u} + (-\mathbf{u})$.

1.8.8 Proposition. Existence d'un sous-espace complémentaire

Supposons que E soit de dimension finie. Pour tout sous-espace vectoriel S de E, il existe un sous-espace vectoriel T de E (non unique) tel que E soit somme directe de S et T. On dit que T est un sous-espace complémentaire *de S dans E.*

DÉMONSTRATION

Ecartons les cas triviaux où $S = \{\mathbf{0}\}$ et $S = E$. Le sous-espace vectoriel S admet alors une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$, où k est inférieur à la dimension n de E. Par le théorème 1.7.3, cette base se prolonge en une base $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{e}_{k+1}, ..., \mathbf{e}_n)$ de E. Soit T le sous-espace vectoriel engendré par la famille $(\mathbf{e}_{k+1}, \mathbf{e}_{k+2}, ..., \mathbf{e}_n)$. Si $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n$ est un vecteur quelconque de E, alors $\mathbf{x} = \mathbf{s} + \mathbf{t}$, où $\mathbf{s} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_k\mathbf{e}_k$ est un vecteur de S et $\mathbf{t} = x_{k+1}\mathbf{e}_{k+1} + x_{k+2}\mathbf{e}_{k+2} + ...$

$+ x_n\mathbf{e}_n$ est un vecteur de T. En outre, $S \cap T = \{\mathbf{0}\}$ car aucun vecteur, excepté le vecteur nul, ne peut être combinaison linéaire des vecteurs $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k$ et des vecteurs $\mathbf{e}_{k+1}, \mathbf{e}_{k+2}, ..., \mathbf{e}_n$. Nous en concluons que $E = S \oplus T$.

Par exemple, si S est un hyperplan, tout vecteur n'appartenant pas à S engendre une droite vectorielle D telle que $E = S \oplus D$.

1.8.9 Proposition. Dimension d'une somme directe

Si E est somme directe de deux sous-espaces vectoriels S et T de dimension finie, alors E est de dimension finie et

$$\dim E = \dim S + \dim T. \tag{1.12}$$

DÉMONSTRATION

Ecartons le cas trivial où un des sous-espaces S et T se réduit à $\{\mathbf{0}\}$. Soit $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$ une base de S et $(\mathbf{e}_{k+1}, \mathbf{e}_{k+2}, ..., \mathbf{e}_n)$ une base de T. Il suffit de démontrer que $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{e}_{k+1}, ..., \mathbf{e}_n)$ est une base de E, ou, ce qui revient au même, que tout vecteur de E s'écrit de manière unique sous la forme d'une combinaison linéaire des vecteurs $\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{e}_{k+1}, ..., \mathbf{e}_n$. Or, cela est immédiat, car tout vecteur de E s'écrit de manière unique sous la forme $\mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T.

1.8.10 Extension

On dit que la somme $S_1 + S_2 + ... + S_k$ des sous-espaces vectoriels $S_1, S_2, ..., S_k$ de E est *directe* si, pour $i = 1, 2, ..., k$, $S_i \cap T_i = \{\mathbf{0}\}$, où T_i est la somme des S_j d'indice j différent de i. On note cette somme directe $S_1 \oplus S_2 \oplus ... \oplus S_k$.

De même qu'en 1.8.7, on démontre qu'une somme $S_1 + S_2 + ... + S_k$ est directe si et seulement si chacun de ses vecteurs s'écrit d'une seule manière sous la forme $\mathbf{s}_1 + \mathbf{s}_2 + ... + \mathbf{s}_k$, où $\mathbf{s}_i$ est un vecteur de S_i pour $i = 1, 2, ..., k$. Il est évident que cette condition est remplie si et seulement si la seule décomposition du vecteur nul est celle dans laquelle tous les $\mathbf{s}_i$ sont nuls.

Si les dimensions de $S_1, S_2, ..., S_k$ sont finies, la relation (1.12) se généralise et devient

$$\dim(S_1 \oplus S_2 \oplus ... \oplus S_k) = \dim S_1 + \dim S_2 + ... + \dim S_k. \tag{1.13}$$

1.9 ESPACES AFFINES

1.9.1 Introduction

L'espace $\mathscr{G}$ de la géométrie élémentaire est à la fois le modèle usuel et la source de la notion d'espace affine que nous allons introduire. Cet espace $\mathscr{G}$ est associé à l'espace vectoriel géométrique V par la correspondance entre flèches et vecteurs étudiée dans la section 1.1. La définition suivante ne fait que mettre en évidence les traits dominants de cette correspondance.

1.9.2 Espaces affines

Soit $\mathscr{E}$ un ensemble non vide d'éléments que nous appellerons *points* et désignerons par les lettres $P, Q, ...$; soit en outre E un espace vectoriel. Supposons qu'à tout couple de points (P, Q) corresponde un vecteur noté $\overrightarrow{PQ}$. On dit que $\mathscr{E}$ est un *espace affine* d'*espace directeur* ou *direction E* si les conditions suivantes sont satisfaites:

(a) Pour tout point P fixé, la correspondance entre couples (P, Q) et vecteurs $\mathbf{x}$ est biunivoque, autrement dit, pour tout vecteur $\mathbf{x}$ il existe un point Q et un seul tel que $\mathbf{x} = \overrightarrow{PQ}$.

(b) Pour tout triplet de points (P, Q, R), $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$ (*relation de Chasles*).

1.9.3 Notation

Si P est un point et $\mathbf{x}$ un vecteur, pour exprimer que Q est l'unique point tel que $\mathbf{x} = \overrightarrow{PQ}$, nous écrirons

$$Q = P + \mathbf{x}. \tag{1.14}$$

Bien qu'un peu abusive, cette écriture est commode à l'usage et suggère bien le sens de l'opération qu'elle désigne (fig. 1.11).

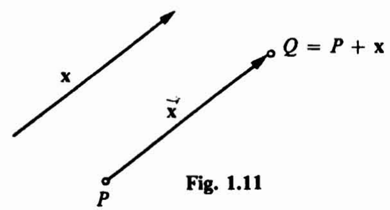

Fig. 1.11

On remarquera que

$$P + (\mathbf{x} + \mathbf{y}) = (P + \mathbf{x}) + \mathbf{y}.$$

1.9.4 Dimension d'un espace affine

On appelle *dimension* d'un espace affine la dimension de son espace directeur.

1.9.5 Règles de calcul dans les espaces affines

Les règles suivantes découlent directement de la définition 1.9.2.

(1) Pour tout point P, $\overrightarrow{PP} = \mathbf{0}$. Cela résulte de la condition (b) appliquée au cas où $P = Q = R$.

(2) Si $\overrightarrow{PQ} = \mathbf{0}$, alors $P = Q$. Cela résulte de la condition (a), compte tenu de la règle (1).

(3) $\overrightarrow{PQ} = -\overrightarrow{QP}$. Il suffit de poser $R = P$ dans la condition (b).

(4) *Règle du parallélogramme.* $\overrightarrow{PP'} = \overrightarrow{QQ'}$ si et seulement si $\overrightarrow{P'Q'} = \overrightarrow{PQ}$. En effet, d'après la condition (b), $\overrightarrow{PP'} + \overrightarrow{P'Q'} = \overrightarrow{PQ'} = \overrightarrow{PQ} + \overrightarrow{QQ'}$ (fig. 1.12).

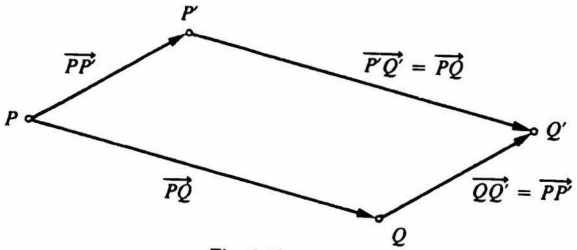

Fig. 1.12

1.9.6 Vectorialisé d'un espace affine

Dans 1.3.2, nous avons montré comment l'espace $\mathscr{G}$ peut être muni d'une structure d'espace vectoriel. Dans le cas général d'un espace affine $\mathscr{E}$, le procédé est le même. On choisit un point quelconque O de $\mathscr{E}$. La correspondance entre couples (O, P) et vecteurs de l'espace directeur E étant alors biunivoque (par (a)), tout comme celle entre couples (O, P) et points P, on définit l'addition de points et la multiplication d'un point par un scalaire par les opérations correspondantes sur les vecteurs de E. Muni de ces deux opérations, $\mathscr{E}$ devient un espace vectoriel, appelé *vectorialisé* de $\mathscr{E}$ relativement à O. Nous désignerons cet espace par $\mathscr{E}_O$ et appellerons O *origine*.

Vu la manière dont les opérations ont été définies, il résulte que $\mathscr{E}_O$ est isomorphe (cf. 1.7.10) à l'espace directeur E:

$$\begin{array}{ccc} \mathscr{E}_O & & E \\ P = O + \mathbf{x} & \longleftrightarrow & \mathbf{x} \end{array}$$

Toutefois, cet isomorphisme dépend du choix de l'origine O et en pratique cette origine est choisie sur la base de données inhérentes aux problèmes posés. Par exemple, si une transformation affine admet un point invariant, nous verrons qu'il y a avantage à choisir ce point comme origine.

1.9.7 Exemples

(1) Il est dit dans 1.9.1 que l'espace $\mathscr{G}$ de la géométrie élémentaire est un espace affine. En effet, sa direction est l'espace géométrique V et les conditions (a) et (b) de la définition 1.9.2 sont satisfaites. Il faut bien noter qu'au couple de points (P, Q) est associé le vecteur $\overrightarrow{PQ}$ et non pas la flèche PQ. En fait, la flèche pouvant être identifée au couple de points, nous voyons que ce que postule la définition 1.9.2 n'est rien d'autre qu'une forme abstraite de correspondance entre flèches et vecteurs.

(2) Tout espace vectoriel E peut être considéré comme un espace affine de direction E lui-même si au couple de vecteurs $(\mathbf{x}, \mathbf{y})$ est associé le vecteur $\mathbf{y} - \mathbf{x}$. En effet, les conditions (a) et (b) de la définition 1.9.2 sont satisfaites.

(3) Soit E un espace vectoriel, S un sous-espace vectoriel de E distinct de E et $\mathbf{x}$ un vecteur de E. Nous désignerons par $\mathbf{x} + S$ l'ensemble des vecteurs $\mathbf{z}$ de la forme $\mathbf{x} + \mathbf{y}$, où $\mathbf{y}$ parcourt S. Si $\mathbf{x}$ n'appartient pas à S, $\mathbf{x} + S$ n'est pas un sous-espace vectoriel de E, car le vecteur nul n'appartient pas à $\mathbf{x} + S$. Par contre, $\mathbf{x} + S$ devient un espace affine de direction S, lorsqu'on y introduit la correspondance entre couples et vecteurs définie dans l'exemple (2). En effet, la différence de deux vecteurs de $\mathbf{x} + S$ est un vecteur de S et les conditions (a) et (b) de la définition 1.9.2 sont satisfaites.

(4) Pour illustrer l'exemple (3), considérons le système d'équations linéaires

$$\begin{aligned} 3x_1 - 2x_2 + 4x_3 &= 3 \\ -x_1 + x_2 - 5x_3 &= -4. \end{aligned} \tag{1.15}$$

Comme dans l'exemple (4) de 1.4.3, nous appellerons solution de ce système tout vecteur-colonne (x_i^0) de $\mathbb{R}^3$ dont les termes x_1^0, x_2^0, x_3^0 vérifient les deux équations. Prenons une solution particulière de ce système, par exemple

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Nous obtenons alors la solution générale en additionnant à cette solution particulière une solution quelconque du système (1.2) (cf. 3.5.5). En d'autres termes, l'ensemble des solutions de (1.15) s'écrit sous la forme

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + S,$$

où S est le sous-espace vectoriel de $\mathbb{R}^3$ formé des solutions du système (1.2).

1.10 SOUS-ESPACES AFFINES, PARALLÉLISME

1.10.1 Introduction

L'exemple (3) de 1.9.7 et son illustration (4) nous suggèrent la notion de sous-espace affine que nous allons introduire. Les sous-espaces affines de l'espace affine $\mathscr{G}$ de la géométrie élémentaire sont les points, les droites, les plans et $\mathscr{G}$ lui-même.

Dans ce qui suit, $\mathscr{E}$ désignera un espace affine de direction E.

1.10.2 Sous-espaces affines

On dit qu'un sous-ensemble $\mathscr{S}$ de $\mathscr{E}$ est un *sous-espace affine* s'il existe un point P_0 de $\mathscr{E}$ et un sous-espace vectoriel S de E tels que

$$\mathscr{S} = \{P: \overrightarrow{P_0P} \in S\} = \{P: P = P_0 + \mathbf{x},\ \mathbf{x} \in S\}. \tag{1.16}$$

En d'autres termes, $\mathscr{S}$ est un sous-espace affine si, pour un point P_0 de $\mathscr{E}$, $\mathscr{S}$ est un sous-espace vectoriel de $\mathscr{E}_{P_0}$ (fig. 1.13).

On notera que $\mathscr{S}$ ainsi défini n'est pas vide, car il comprend au moins le point P_0.

Suivant l'exemple (3) de 1.9.7, nous désignerons le dernier membre de (1.16) plus brièvement par $P_0 + S$, ce qui nous permettra de considérer un sous-espace affine comme un sous-ensemble $\mathscr{S}$ de $\mathscr{E}$ de la forme

$$\mathscr{S} = P_0 + S, \tag{1.17}$$

où P_0 est un point de $\mathscr{E}$ et S un sous-espace vectoriel de E.

On prendra garde à bien distinguer les différentes significations du signe +: addition dans E, «addition» d'un point et d'un vecteur, «addition» d'un point et d'un sous-espace vectoriel.

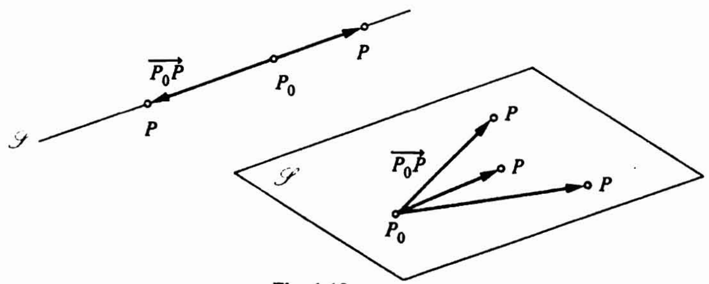

Fig. 1.13

1.10.3 Proposition

Soit $\mathscr{S}$ le sous-espace affine défini par P_0 et S.

(a) *Si P est un point quelconque de $\mathscr{S}$, alors $\mathscr{S} = P + S$.*

(b) $S = \{\mathbf{x}: \mathbf{x} = \overrightarrow{PQ},\ P, Q \in \mathscr{S}\}$.

DÉMONSTRATION

Assertion (a). Posons $\mathscr{S}' = P + S$. Il suffit de montrer que $\mathscr{S}'$ est inclus dans $\mathscr{S}$, car l'argument symétrique nous fournira l'inclusion opposée et donc la conclusion $\mathscr{S} = \mathscr{S}'$. Soit Q un point de $\mathscr{S}'$, c'est-à-dire un point tel que $\overrightarrow{PQ}$ appartienne à S. Puisque P est un point de $\mathscr{S}$, $\overrightarrow{P_0P}$ appartient à S. Mais $\overrightarrow{P_0Q} = \overrightarrow{P_0P} + \overrightarrow{PQ}$, par la condition (b) de la définition 1.9.2, donc $\overrightarrow{P_0Q}$ appartient aussi à S et, par conséquent, Q est un point de $\mathscr{S}$.

Assertion (b). En vertu de (a), pour tout couple (P, Q) de points de $\mathscr{S}$, $\overrightarrow{PQ}$ est un vecteur de S. Réciproquement, si $\mathbf{x}$ est un vecteur de S, choisissons un point quelconque P de $\mathscr{S}$ et posons $Q = P + \mathbf{x}$. D'après (a), Q est un point de $\mathscr{S}$, donc $\mathbf{x}$ est bien un vecteur de la forme $\overrightarrow{PQ}$, avec P et Q des points de $\mathscr{S}$.

1.10.4 Espace directeur

D'après la proposition 1.10.3, le sous-espace affine $\mathscr{S}$ défini par (1.16) ne dépend pas du choix de l'«origine» P_0 et détermine le sous-espace vectoriel S. On dit que S est l'*espace directeur* ou la *direction* de $\mathscr{S}$.

1.10.5 Sous-espaces affines comme espaces affines

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S. A l'aide de la proposition 1.10.3, nous voyons que la correspondance héritée de $\mathscr{E}$ entre couples de points de $\mathscr{S}$ et vecteurs de S fait de $\mathscr{S}$ un espace affine dans le sens de la définition 1.9.2.

1.10.6 Dimension d'un sous-espace affine

On appelle *dimension* d'un sous-espace affine $\mathscr{S}$ de $\mathscr{E}$ la dimension de $\mathscr{S}$ en tant qu'espace affine, c'est-à-dire la dimension de l'espace directeur de $\mathscr{S}$.

1.10.7 Intersection de sous-espaces affines

Soit $\mathscr{S}$ et $\mathscr{T}$ deux sous-espaces affines de $\mathscr{E}$. Si $\mathscr{S}$ et $\mathscr{T}$ ont au moins un point commun P, leur intersection est un sous-espace affine de $\mathscr{E}$ de direction l'intersection de leurs directions. En effet, $\mathscr{S} \cap \mathscr{T} = \{Q: \overrightarrow{PQ} \in S\} \cap \{Q: \overrightarrow{PQ} \in T\} = \{Q: \overrightarrow{PQ} \in S \cap T\}$, où S et T sont les directions respectives de $\mathscr{S}$ et de $\mathscr{T}$.

1.10.8 Sous-espaces affines engendrés

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et $P_1, P_2, ..., P_l$ des points de $\mathscr{E}$. On appelle sous-espace affine *engendré* par $\mathscr{S}$ et $P_1, P_2, ..., P_l$ le plus petit sous-espace affine $\mathscr{T}$ contenant $\mathscr{S}$ et comprenant $P_1, P_2, ..., P_l$. Si S admet une base $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$, on voit aisément que la direction de $\mathscr{T}$ est le sous-espace vectoriel engendré par les vecteurs $\mathbf{v}_1, ..., \mathbf{v}_k, \overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_l}$, où P_0 est un point quelconque de $\mathscr{S}$. Par exemple, le sous-espace affine de $\mathscr{G}$ engendré par une droite et un point est le plan passant par cette droite et ce point (supposé hors de la droite).

1.10.9 Exemples

(1) Comme déjà annoncé et illustré (fig. 1.13), les sous-espaces affines de $\mathscr{G}$ sont les points, les droites, les plans et $\mathscr{G}$ lui-même.

(2) L'espace affine $\mathscr{E}$ est un sous-espace affine de lui-même. Pour s'en rendre compte, il suffit de mettre E à la place de S dans (1.16).

(3) Pour tout point P_0 de $\mathscr{E}$, $\{P_0\}$ est un sous-espace affine. On voit cela en posant $S = \{\mathbf{0}\}$ dans (1.16).

(4) Un sous-espace affine de dimension nulle se réduit à un seul point. Un sous-espace affine de dimension 1 est appelé *droite affine* ou simplement *droite*; un de dimension 2 *plan affine* ou simplement *plan*. Il est clair qu'une droite est déterminée par deux de ses points et un plan par trois de ses points non alignés, c'est-à-dire n'appartenant pas à une même droite.

1.10.10 Hyperplans affines

Si $\mathscr{E}$ est de dimension finie non nulle n, un sous-espace affine de $\mathscr{E}$ de dimension $n - 1$ est appelé *hyperplan affine* ou simplement *hyperplan*. Un hyperplan est donc un sous-espace affine de direction un hyperplan vectoriel. Par exemple, un hyperplan est un point si $n = 1$, une droite si $n = 2$ et un plan si $n = 3$.

1.10.11 Parallélisme

Soit $\mathscr{S}$ et $\mathscr{T}$ des sous-espaces affines de $\mathscr{E}$ de directions respectives S et T. On dit que $\mathscr{S}$ et $\mathscr{T}$ sont *parallèles* si l'une des directions S ou T est incluse dans l'autre. Si $S = T$, on dit aussi que $\mathscr{S}$ et $\mathscr{T}$ sont *parallèles au sens étroit*.

On remarquera que ces deux notions de parallélisme s'équivalent lorsqu'elles sont appliquées à des sous-espaces affines de même dimension finie (en raison de la seconde partie de la proposition 1.8.5). En les appliquant au cas particulier de $\mathscr{G}$, nous retrouvons les notions usuelles de parallélisme entre droites, entre plans et entre droites et plans.

1.10.12 Deux résultats

Formulées en termes de droites et de plans de $\mathscr{G}$, les deux assertions suivantes deviennent des énoncés bien connus de la géométrie élémentaire.

(1) *Généralisation du cinquième postulat d'Euclide.* Soit P un point de $\mathscr{E}$ et $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$. Il existe un unique sous-espace affine $\mathscr{T}$ de $\mathscr{E}$ parallèle à $\mathscr{S}$ au sens étroit et comprenant P. En effet, ce sous-espace affine est tout simplement $\mathscr{T} = \{Q: \overrightarrow{PQ} \in S\}$, où S est la direction de $\mathscr{S}$.

(2) *Position relative de deux sous-espaces affines parallèles.* Si deux sous-espaces affines sont parallèles, soit ils sont disjoints, soit l'un d'entre eux est inclus dans l'autre. S'ils sont parallèles au sens étroit, soit ils sont disjoints, soit confondus. Il suffit en effet de choisir comme point P_0 des représentations (1.16) des deux sous-espaces affines un point de leur intersection (dans le cas où celle-ci n'est pas vide), pour que la conclusion découle de la définition 1.10.11.

1.11 REPÈRES, REPRÉSENTATION PARAMÉTRIQUE, GÉOMÉTRIE ANALYTIQUE AFFINE

1.11.1 Introduction

Comme pour le choix d'une base dans un espace vectoriel, le choix d'un repère dans un espace affine de dimension n permettra d'identifier cet espace à $\mathbb{R}^n$ et, par ce moyen, de traiter les problèmes géométriques par des calculs sur les coordonnées (géométrie analytique).

Dans cette section, $\mathscr{E}$ désignera un espace affine de dimension finie non nulle n et de direction E.

1.11.2 Repères

On appelle *repère* de $\mathscr{E}$ tout ensemble $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ formé d'un point O, appelé *origine*, et d'une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E.

1.11.3 Remarque

On peut concevoir un repère sous la forme d'une famille de points $(O; P_1, P_2, ..., P_n)$ telle que $(\overrightarrow{OP_1}, \overrightarrow{OP_2}, ..., \overrightarrow{OP_n})$ est une base de E.

1.11.4 Coordonnées

On appelle *coordonnées cartésiennes* ou simplement *coordonnées* d'un point P dans un repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ les composantes du vecteur $\overrightarrow{OP}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

Si $x_1, x_2, ..., x_n$ sont les coordonnées d'un point P et $y_1, y_2, ..., y_n$ celles d'un point Q, les composantes du vecteur $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$ sont $y_1 - x_1, y_2 - x_2, ..., y_n - x_n$. Inversement, si $x_1, x_2, ..., x_n$ sont les coordonnées d'un point P et $z_1, z_2, ..., z_n$ les composantes d'un vecteur $\mathbf{z}$, les coordonnées du point $Q = P + \mathbf{z}$ sont $x_1 + z_1, x_2 + z_2, ..., x_n + z_n$, car P et Q sont liés par la relation $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \mathbf{z}$.

Par la suite, un point P défini par ses coordonnées $x_1, x_2, ..., x_n$ sera désigné par $P(x_1, x_2, ..., x_n)$.

1.11.5 Représentation paramétrique d'un sous-espace affine

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et de dimension non nulle k. Soit en outre P_0 un point de $\mathscr{S}$ et $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ une base de S. D'après (a) de la proposition 1.10.3, un point P appartient à $\mathscr{S}$ si et seulement si $P = P_0 + \mathbf{x}$, où $\mathbf{x}$ est un vecteur de S, c'est-à-dire un vecteur de la forme $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k$. Il s'ensuit que $\mathscr{S}$ est l'ensemble des points P satisfaisant à la relation

$$P = P_0 + \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k, \tag{1.18}$$

où $\alpha_1, \alpha_2, ..., \alpha_k$ sont des variables parcourant $\mathbb{R}$. Cette relation est appelée *représentation paramétrique* de $\mathscr{S}$. Les vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ sont appelés *vecteurs directeurs* de $\mathscr{S}$ et les variables $\alpha_1, \alpha_2, ..., \alpha_k$ *paramètres* de la représentation. On notera que le nombre de paramètres est égal à la dimension de $\mathscr{S}$ (fig. 1.14).

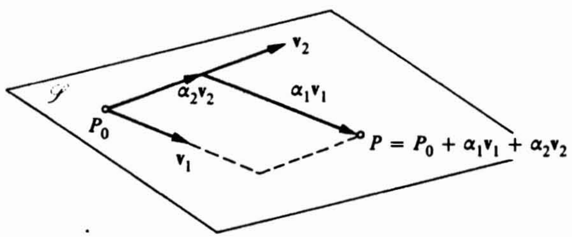

Fig. 1.14

1.11.6 Equations paramétriques

Supposons maintenant que $\mathscr{E}$ soit muni d'un repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. En désignant par $x_1, x_2, ..., x_n$ et $x_1^0, x_2^0, ..., x_n^0$ les coordonnées respectives de P et P_0 et par $v_{1i}, v_{2i}, ..., v_{ni}$ les composantes de $\mathbf{v}_i$, nous voyons que la relation (1.18) s'écrit, de manière équivalente, sous la forme

$$\begin{aligned} x_1 &= x_1^0 + \alpha_1 v_{11} + \alpha_2 v_{12} + ... + \alpha_k v_{1k} \\ x_2 &= x_2^0 + \alpha_1 v_{21} + \alpha_2 v_{22} + ... + \alpha_k v_{2k} \\ &...... \\ x_n &= x_n^0 + \alpha_1 v_{n1} + \alpha_2 v_{n2} + ... + \alpha_k v_{nk}. \end{aligned} \tag{1.19}$$

Ces équations sont appelées *équations paramétriques* de $\mathscr{S}$.

1.11.7 Cas particuliers

Une droite est déterminée par un de ses points et un vecteur directeur, un plan par un de ses points et deux vecteurs directeurs.

	Droite:	Plan:
Représentation paramétrique:	$P = P_0 + \alpha\mathbf{v}$	$P = P_0 + \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$
Equations paramétriques dans le cas $n = 3$:	$x_1 = x_1^0 + \alpha v_1$ $x_2 = x_2^0 + \alpha v_2$ $x_3 = x_3^0 + \alpha v_3$	$x_1 = x_1^0 + \alpha_1 v_{11} + \alpha_2 v_{12}$ $x_2 = x_2^0 + \alpha_1 v_{21} + \alpha_2 v_{22}$ $x_3 = x_3^0 + \alpha_1 v_{31} + \alpha_2 v_{32}$.

(1.20)

1.11.8 Equation cartésienne d'un hyperplan

Supposons que n soit supérieur à 1. Si $k = n - 1$, les équations paramétriques (1.19) sont celles d'un hyperplan. Par le procédé d'élimination (cf. exemple (2) de 3.1.6), $n - 1$ équations peuvent être utilisées pour éliminer les $n - 1$ paramètres $\alpha_1, \alpha_2, ..., \alpha_{n-1}$ de l'équation restante. Le résultat sera une relation linéaire entre les coordonnées $x_1, x_2, ..., x_n$, plus exactement une équation de la forme

$$\nu_1 x_1 + \nu_2 x_2 + ... + \nu_n x_n = \delta, \tag{1.21}$$

où $\nu_1, \nu_2, ..., \nu_n$ sont des nombres non tous nuls et δ est un nombre pouvant s'annuler, auquel cas l'hyperplan passe par l'origine O. Cette équation est appelée *équation cartésienne* ou simplement *équation de l'hyperplan*. Les coordonnées d'un point P satisfont à (1.21) si et seulement si P appartient à l'hyperplan.

Réciproquement, toute équation de la forme (1.21) est l'équation d'un hyperplan, à condition, bien entendu, que les coefficients $\nu_1, \nu_2, ..., \nu_n$ ne soient pas tous nuls. En effet, si par exemple ν_n est non nul, les solutions de (1.21) peuvent être écrites sous la forme

$$\begin{aligned} x_1 &= \alpha_1 \\ x_2 &= \alpha_2 \\ &...... \\ x_{n-1} &= \alpha_{n-1} \\ x_n &= \frac{\delta}{\nu_n} - \frac{\nu_1}{\nu_n}\alpha_1 - ... - \frac{\nu_{n-1}}{\nu_n}\alpha_{n-1}, \end{aligned}$$

où $\alpha_1, \alpha_2, ..., \alpha_{n-1}$ sont des variables parcourant $\mathbb{R}$. Ces équations sont de la forme (1.19) et représentent donc un hyperplan.

Dans le cas particulier où $n = 2$,

$$\nu_1 x_1 + \nu_2 x_2 = \delta$$

est l'équation cartésienne d'une droite et dans celui où $n = 3$,

$$\nu_1 x_1 + \nu_2 x_2 + \nu_3 x_3 = \delta$$

est l'équation cartésienne d'un plan.

1.11.9 Problèmes de géométrie analytique affine

La géométrie analytique affine traite les problèmes de la géométrie affine (parallélisme, incidence, ...) par des calculs dans $\mathbb{R}^n$. Voici quelques problèmes résolus.

1 (1) Trouver les équations paramétriques du sous-espace affine engendré par les points $P_0(1, 0, 3, -1)$, $P_1(-1, 2, 0, 3)$, $P_2(0, 0, -1, 2)$, $P_3(-3, 5, 1, 2)$.

Solution. Trois vecteurs directeurs de ce sous-espace affine sont

$$\begin{aligned} \mathbf{v}_1 &= \overrightarrow{P_0P_1}(-2, 2, -3, 4) \\ \mathbf{v}_2 &= \overrightarrow{P_0P_2}(-1, 0, -4, 3) \\ \mathbf{v}_3 &= \overrightarrow{P_0P_3}(-4, 5, -2, 3)\,. \end{aligned}$$

Ses équations paramétriques sont donc

$$\begin{aligned} x_1 &= 1 - 2\alpha_1 - \alpha_2 - 4\alpha_3 \\ x_2 &= 2\alpha_1 + 5\alpha_3 \\ x_3 &= 3 - 3\alpha_1 - 4\alpha_2 - 2\alpha_3 \\ x_4 &= -1 + 4\alpha_1 + 3\alpha_2 + 3\alpha_3\,. \end{aligned}$$

2 (2) Trouver l'intersection des deux plans d'équations paramétriques

$$\begin{aligned} x_1 &= 1 + \alpha_1 + 2\alpha_2 & x_1 &= 1 - \alpha'_1 \\ x_2 &= 1 - \alpha_1 + \alpha_2 & x_2 &= 1 + \alpha'_1 + \alpha'_2 \\ x_3 &= -1 + 2\alpha_1 - \alpha_2 & x_3 &= -1 + 2\alpha'_2 \\ x_4 &= -2 + \alpha_1 + \alpha_2\,, & x_4 &= -2 + 2\alpha'_1 + 3\alpha'_2\,. \end{aligned}$$

Solution. Deux vecteurs directeurs du premier plan sont $\mathbf{v}_1(1, -1, 2, 1)$ et $\mathbf{v}_2(2, 1, -1, 1)$, deux du deuxième $\mathbf{v}'_1(-1, 1, 0, 2)$ et $\mathbf{v}'_2(0, 1, 2, 3)$. Ces quatre vecteurs sont linéairement indépendants, donc aucun vecteur, sauf le vecteur nul, ne peut être en même temps combinaison linéaire de $\mathbf{v}_1$ et $\mathbf{v}_2$ et de $\mathbf{v}'_1$ et $\mathbf{v}'_2$. Cela entraîne que l'intersection des deux espaces directeurs est $\{\mathbf{0}\}$. D'autre part, en posant $\alpha_1 = \alpha_2 = \alpha'_1 = \alpha'_2 = 0$ dans les équations paramétriques, nous voyons que le point $P_0(1, 1, -1, -2)$ appartient aux deux plans. Nous en concluons que l'intersection cherchée se réduit au point P_0.

3 (3) Deux droites $\mathscr{D}$ et $\mathscr{D}'$ sont données par leurs équations paramétriques

$$\begin{aligned} x_1 &= -\tfrac{2}{3} - 2\alpha & x_1 &= \alpha' \\ x_2 &= \tfrac{3}{2} + 3\alpha & x_2 &= -1 + 2\alpha' \\ x_3 &= 6\alpha\,, & x_3 &= 4 + \alpha'\,. \end{aligned}$$

Déterminer les équations paramétriques de la droite passant par le point $P_0(1, 1, -1)$ et rencontrant $\mathscr{D}$ et $\mathscr{D}'$.

Solution. Les points d'intersection P de $\mathscr{D}$ et la droite cherchée et P' de $\mathscr{D}'$ et cette même droite satisfont à la relation $\overrightarrow{P_0P} = \beta\overrightarrow{P_0P'}$, où β est un nombre inconnu non nul. En composantes cette relation s'exprime par les trois équations

$$\begin{aligned} -\tfrac{5}{3} - 2\alpha &= \beta(-1 + \alpha') \\ \tfrac{1}{2} + 3\alpha &= \beta(-2 + 2\alpha') \\ 1 + 6\alpha &= \beta(5 + \alpha')\,. \end{aligned}$$

En multipliant la deuxième équation par 2 et en la soustrayant de la troisième, nous obtenons aussitôt $\alpha' = 3$, ce qui nous permet de calculer les coordonnées de P', à savoir 3, 5, 7. En prenant alors pour vecteur directeur de la droite cherchée le vecteur $\frac{1}{2}\overrightarrow{P_0P'}$, nous voyons que les équations paramétriques de celle-ci s'écrivent

$$\begin{aligned} x_1 &= 1 + \alpha \\ x_2 &= 1 + 2\alpha \\ x_3 &= -1 + 4\alpha\,. \end{aligned}$$

Une autre manière de résoudre ce problème consiste à calculer l'intersection des deux plans définis par P_0 et chacune des deux droites $\mathscr{D}$ et $\mathscr{D}'$.

4 (4) Déterminer la projection du point $P(1, -3, -2, 1)$ sur l'hyperplan d'équation

$$x_1 - x_2 + 3x_3 - 4x_4 = 4,$$

parallèlement à une droite de vecteur directeur $\mathbf{v}(1, 4, 3, -1)$.

Solution. Les équations paramétriques de la droite passant par P de vecteur directeur $\mathbf{v}$ sont

$$\begin{aligned} x_1 &= 1 + \alpha \\ x_2 &= -3 + 4\alpha \\ x_3 &= -2 + 3\alpha \\ x_4 &= 1 - \alpha\,. \end{aligned}$$

La projection cherchée est le point d'intersection de cette droite et l'hyperplan donné. Ses coordonnées s'obtiennent en déterminant la valeur de α pour laquelle les seconds membres des équations paramétriques vérifient l'équation cartésienne de l'hyperplan. Cette valeur est $\alpha = 1$, donc les coordonnées de la projection de P sont 2, 1, 1, 0.

1.12 EXERCICES

1.12.1 Montrer que les vecteurs-colonnes

$$\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}-1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}2\\1\\1\end{pmatrix}$$

engendrent $\mathbb{R}^3$ et exprimer le vecteur-colonne $\begin{bmatrix} -3 \\ 4 \\ -1 \end{bmatrix}$ comme combinaison linéaire de ces vecteurs

1.12.2 Montrer que toute fonction de la forme $\mathbf{f}(t) \equiv \alpha \sin(\lambda t + \mu)$ $(\alpha, \lambda \geqslant 0, \mu \in [0, 2\pi))$ est combinaison linéaire des fonctions $\mathbf{s}(t) \equiv \sin(\lambda t)$ et $\mathbf{c}(t) \equiv \cos(\lambda t)$ et, inversement, que toute combinaison linéaire de $\mathbf{s}$ et $\mathbf{c}$ est une fonction de la même forme que $\mathbf{f}$.

1.12.3 Exprimer le polynôme $\mathbf{p}(t) \equiv 1 + t^4$ sous la forme d'une combinaison linéaire des polynômes $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ et $\mathbf{p}_4$ définis dans (1.5).

1.12.4 Démontrer que les vecteurs-colonnes

$$\begin{bmatrix} a_1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ 0 \end{bmatrix}, \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}$$

sont linéairement indépendants si et seulement si a_1, b_2, c_3 et d_4 sont différents de 0.

1.12.5 En dérivant n fois la relation $\alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n \equiv 0$, montrer que les monômes $\mathbf{p}_0$, $\mathbf{p}_1$, ..., $\mathbf{p}_n$ définis dans (1.8) sont linéairement indépendants.

1.12.6 Montrer que les fonctions $\mathbf{f}_1(t) \equiv 1$, $\mathbf{f}_2(t) \equiv e^t$ et $\mathbf{f}_3(t) \equiv e^{2t}$ sont linéairement indépendantes.

$: \left(\sum_{i=0}^{2} \alpha_i e^{it}\right| , \forall t, \forall i, e^{it} \geq 0 \Rightarrow \alpha_i \geq 0.$

1.12.7 Soit $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ une famille libre de vecteurs d'un espace vectoriel E. Soit $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k)$ une deuxième famille de vecteurs de E. On suppose que chaque vecteur $\mathbf{x}_i$ soit combinaison linéaire des vecteurs $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k$. Montrer que la famille $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k)$ est libre.

1.12.8 Soit $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ une base d'un espace vectoriel E de dimension 4.
(a) Montrer que les vecteurs $\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_4$, $\mathbf{v}_2 = \mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3 - 2\mathbf{e}_4$ et $\mathbf{v}_3 = -\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 + \mathbf{e}_4$ sont linéairement indépendants.
(b) Prolonger la famille $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ en une base de E.

1.12.9 Soit $(\mathbf{e}_1, \mathbf{e}_2)$ une base d'un espace vectoriel E de dimension 2. Montrer que les familles de vecteurs $(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_2)$, $(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2)$ et $(\mathbf{e}_1, \mathbf{e}_2 + \alpha\mathbf{e}_1)$ (α étant un nombre arbitraire) sont des bases de E. Calculer, en outre, les composantes de $\mathbf{e}_1$ et $\mathbf{e}_2$ dans chacune de ces bases.

1.12.10 Dans chacun des cas suivants, dire si la famille $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ est une base de l'espace vectoriel des polynômes de degré inférieur ou égal à 2. Si oui, trouver les composantes dans cette base des monômes $\mathbf{p}_0(t) \equiv 1$, $\mathbf{p}_1(t) \equiv t$ et $\mathbf{p}_2(t) \equiv t^2$.
(a) $\mathbf{f}_1(t) \equiv 1 + t^2$, $\mathbf{f}_2(t) \equiv 2 - t + t^2$, $\mathbf{f}_3(t) \equiv -6 + 3t - 3t^2$.
(b) $\mathbf{f}_1(t) \equiv 1 + 2t + t^2$, $\mathbf{f}_2(t) \equiv 1 - 2t + t^2$, $\mathbf{f}_3(t) \equiv t$.
(c) $\mathbf{f}_1(t) \equiv t$, $\mathbf{f}_2(t) \equiv t + t^2$, $\mathbf{f}_3(t) \equiv 1 + t + t^2$.

1.12.11 Les sous-ensembles suivants sont-ils des sous-espaces vectoriels de $\mathbb{R}^5$?

$\{(a_i): a_1 = 0\}$, $\{(a_i): a_2 = 1\}$, $\{(a_i): 2a_1 - a_4 = 0\}$, $\{(a_i): a_1 a_4 = a_5\}$, $\{(a_i): a_1 \text{ rationnel}\}$.

1.12.12 Soit $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_6)$ une base d'un espace vectoriel E de dimension 6. Soit S le sous-espace vectoriel de E engendré par les vecteurs $\mathbf{e}_1 - \mathbf{e}_2$, $\mathbf{e}_3 + \mathbf{e}_4$ et $\mathbf{e}_5 + \mathbf{e}_6$. Quelles conditions doivent satisfaire les composantes d'un vecteur $\mathbf{x}$ de E pour qu'il appartienne à S?

1.12.13 Soit E un espace vectoriel de dimension 4, muni d'une base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$.
(a) Montrer que l'ensemble S des vecteurs $\mathbf{x}$ de E dont les composantes x_1, x_2, x_3, x_4 satisfont à la condition $x_1 - 2x_2 + x_3 - x_4 = 0$ est un hyperplan vectoriel.
(b) Exhiber une base de S.
(c) Trouver un sous-espace complémentaire de S dans E.

1.12.14 Soit S et T deux sous-espaces vectoriels d'un espace vectoriel E.
(a) Démontrer que la somme $S + T$ et l'intersection $S \cap T$ sont des sous-espaces vectoriels de E.
(b) Démontrer que $S + T$ est l'intersection de tous les sous-espaces vectoriels U de E tels que $U \supset S \cup T$.

1.12.15 Soit S et T les sous-espaces vectoriels de $\mathbb{R}^4$ engendrés par les familles respectives

$$\left(\begin{bmatrix}1\\-1\\2\\1\end{bmatrix}, \begin{bmatrix}2\\0\\-1\\0\end{bmatrix}, \begin{bmatrix}0\\-2\\5\\-1\end{bmatrix}\right) \text{ et } \left(\begin{bmatrix}1\\-1\\2\\0\end{bmatrix}, \begin{bmatrix}4\\-2\\3\\1\end{bmatrix}, \begin{bmatrix}3\\-1\\1\\-1\end{bmatrix}\right).$$

Déterminer les dimensions de S et de T, ainsi que celles de $S + T$ et de $S \cap T$.

1.12.16 Soit S, T et U trois sous-espaces vectoriels d'un espace vectoriel E.

(a) Montrer que $S + (T \cap U) \subset (S + T) \cap (S + U)$.

(b) Montrer que si $S \subset T$, les deux membres de l'inclusion sont égaux.

(c) Trouver un exemple montrant que l'égalité n'est généralement pas vraie.

1.12.17 Soit S et T deux sous-espaces vectoriels de dimension finie d'un espace vectoriel E. Montrer que

$$\dim(S + T) + \dim(S \cap T) = \dim S + \dim T.$$

En déduire la relation (1.12).

Exercices sur les espaces affines

1.12.18 Soit $\mathcal{S}$ l'hyperplan d'équation $3x_1 - x_2 + x_3 - 2x_4 = 1$ et $\mathcal{D}$ le lieu géométrique défini par les équations $\dfrac{x_1 - 1}{-2} = \dfrac{x_2 - 1}{2} = x_3 - 2 = \dfrac{x_4 + 1}{-4}$.

(a) Montrer que $\mathcal{D}$ est une droite et écrire ses équations paramétriques.

(b) Calculer les coordonnées du point d'intersection de $\mathcal{D}$ et $\mathcal{S}$.

1.12.19 Déterminer la projection parallèlement à une droite de vecteur directeur $\mathbf{v}(-1, 2, 1, 2, -2)$ de la droite d'équations $x_1 - 2 = \dfrac{x_2 + 1}{2} = \dfrac{x_3 - 5}{3} = \dfrac{x_4 + 2}{-1} = \dfrac{x_5 - 1}{4}$ sur l'hyperplan d'équation $2x_1 + 3x_2 - x_5 = 2$.

1.12.20 Dans chacun des cas suivants, étudier la position relative des sous-espaces affines $\mathcal{S}$ et $\mathcal{T}$.

(a) $\mathcal{S}$ est l'hyperplan d'équation $x_1 - x_2 + x_4 = 1$ et $\mathcal{T}$ le plan passant par les points $P_1(1, 1, 2, 2)$, $P_2(2, 2, 4, 2)$ et $P_3(1, 3, 1, 4)$.

(b) $\mathcal{S}$ et $\mathcal{T}$ sont les plans d'équations paramétriques respectives

$$\begin{aligned} x_1 &= 1 + \alpha_1 - \alpha_2 \\ x_2 &= 2 + \alpha_1 + \alpha_2 \\ x_3 &= -3 + \alpha_1 - \alpha_2 \\ x_4 &= 1 + \alpha_1 + \alpha_2, \end{aligned} \qquad \begin{aligned} x_1 &= 2 - \alpha_1' - \alpha_2' \\ x_2 &= 3 + \alpha_1' - \alpha_2' \\ x_3 &= -2 + \alpha_1' + \alpha_2' \\ x_4 &= 4 - \alpha_1' + \alpha_2'. \end{aligned}$$

(c) $\mathcal{S}$ est l'hyperplan d'équation $3x_1 - x_3 + 2x_4 = 2$ et $\mathcal{T}$ la droite passant par le point $P_0(1, 5, 3, 1)$ et de vecteur directeur $\mathbf{v}(1, 0, 1, -1)$.

(d) $\mathcal{S}$ est l'hyperplan d'équation $x_1 - 3x_2 + x_4 = 3$ et $\mathcal{T}$ le plan passant par le point $P_0(1, 0, 3, -2)$ et de vecteurs directeurs $\mathbf{v}_1(-1, 0, 1, 1)$ et $\mathbf{v}_2(1, 1, -2, 1)$.

(e) $\mathcal{S}$ est l'hyperplan d'équation $x_1 - 3x_2 + x_3 - 2x_4 = 4$ et $\mathcal{T}$ l'hyperplan d'équations paramétriques

$$\begin{aligned} x_1 &= 1 + \alpha_1 - 2\alpha_2 + 2\alpha_3 \\ x_2 &= 1 + \alpha_1 + \alpha_2 + \alpha_3 \\ x_3 &= 2 - \alpha_2 - \alpha_3 \\ x_4 &= -2 - \alpha_1 - 3\alpha_2 - \alpha_3. \end{aligned}$$

1.12.21 Soit $\mathscr{E}_O$ le vectorialisé relativement à une origine O d'un espace affine $\mathscr{E}$ de dimension finie supérieure à 1 ou infinie. Etudier le lieu géométrique des points de la forme $(\alpha - \beta)P + \beta Q$, où P et Q sont des points distincts de $\mathscr{E}_O$, $\alpha \in (-\infty, 1]$ et $\beta \in \mathbb{R}$.

1.12.22 Soit $\mathscr{E}_O$ le vectorialisé relativement à une origine O d'un espace affine $\mathscr{E}$ de dimension 3. Soit P_1, P_2, P_3 et P_4 quatre points n'appartenant pas à un même plan. Montrer que l'ensemble des combinaisons convexes de P_1, P_2, P_3 et P_4 est le tétraèdre dont les sommets sont ces quatre points.

1.12.23 Soit $\mathscr{E}$ un espace affine muni d'une origine O. Soit $(P_1, P_2, ..., P_k)$ une famille de points de $\mathscr{E}$ et $(\alpha_1, \alpha_2, ..., \alpha_k)$ une famille de nombres telle que $\alpha = \alpha_1 + \alpha_2 + ... + \alpha_k \neq 0$. On appelle le point G de $\mathscr{E}$ défini par la relation $\overrightarrow{OG} = \dfrac{1}{\alpha}(\alpha_1\overrightarrow{OP_1} + \alpha_2\overrightarrow{OP_2} + ... + \alpha_k\overrightarrow{OP_k})$ *barycentre* des points $P_1, P_2, ..., P_k$ affectés des coefficients respectifs $\alpha_1, \alpha_2, ..., \alpha_k$. Lorsque $\alpha_1 = \alpha_2 = ... = \alpha_k = \dfrac{1}{k}$, on appelle G *centre de gravité* des points $P_1, P_2, ..., P_k$.

(a) Montrer que la définition de G ne dépend pas du choix de l'origine O.

(b) Situer le centre de gravité des sommets d'un triangle.

1.12.24 Suite de l'exercice précédent. Soit $N_1, N_2, \ldots, N_l$ les ensembles d'une partition de $\{1, 2, \ldots, k\}$. Pour $i = 1, 2, \ldots, l$, on suppose que $\beta_i = \sum_{j \in N_i} \alpha_j$ soit non nul et on désigne par G_i le barycentre des points P_j avec $j \in N_i$, affectés des coefficients α_j.

(a) Montrer que G (défini dans l'exercice précédent) est le barycentre des points $G_1, G_2, \ldots, G_l$ affectés des coefficients respectifs $\beta_1, \beta_2, \ldots, \beta_l$.

(b) A l'aide de (a), situer le centre de gravité des sommets d'un tétraèdre.

1.12.25 Soit $\mathscr{S}$, $\mathscr{S}'$ et $\mathscr{S}''$ trois hyperplans parallèles d'un espace affine de dimension finie supérieure à 1. Soit en outre $\mathscr{D}_1, \mathscr{D}_2, \ldots$ des droites du même espace, non parallèles à $\mathscr{S}$. On désigne par P_i, P_i' et P_i'' les points d'intersection respectifs de $\mathscr{S}$, $\mathscr{S}'$, $\mathscr{S}''$ et $\mathscr{D}_i$. Montrer que $\overrightarrow{P_iP_i''} = \alpha\overrightarrow{P_iP_i'}$, avec α indépendant de i (*théorème de Thalès*).

CHAPITRE 2

Espaces vectoriels euclidiens et espaces affines euclidiens

2.1 PRODUIT SCALAIRE DANS L'ESPACE VECTORIEL GÉOMÉTRIQUE

2.1.1 Introduction

Dans le présent chapitre, nous étudierons les notions qui dérivent de la donnée d'une nouvelle opération sur les vecteurs, celle de produit scalaire. Il s'agit de notions métriques telles que la norme ou l'orthogonalité. Dans l'espace vectoriel géométrique V, on peut définir un produit scalaire en partant des idées de longueur et d'angle. Muni de ce produit scalaire, V devient un exemple de ce que nous appellerons *espace vectoriel euclidien* (cf. définition 2.2.3).

Tout au long de cette section, les vecteurs seront les éléments de V.

2.1.2 Longueur, angle

En choisissant une unité de longueur, nous pouvons mesurer l'intensité de chaque flèche, autrement dit, déterminer sa *longueur*. Nous pouvons aussi mesurer l'écart angulaire de deux flèches quelconques d'origine commune (non nécessairement distinctes) en prenant comme unité de mesure par exemple le radian. La mesure de cet écart est alors un nombre compris entre 0 et π, appelé *angle* des deux flèches. Si les deux flèches ont même direction et même sens, leur angle est nul et si elles ont même direction et sens opposé, ce même angle est π.

2.1.3 Norme, angle de deux vecteurs

Les flèches représentatives d'un même vecteur $\mathbf{x}$ ont toutes la même longueur. Nous désignerons cette longueur par $\|\mathbf{x}\|$ et l'appellerons *norme* de $\mathbf{x}$. Il est clair qu'un vecteur est nul si et seulement si sa norme est nulle. Nous dirons qu'un vecteur est *unitaire* si sa norme est 1. Si $\mathbf{x}$ est un vecteur non nul, $\frac{1}{\|\mathbf{x}\|}\mathbf{x}$ est un vecteur unitaire, puisque la norme de $\alpha\mathbf{x}$ est $|\alpha|\,\|\mathbf{x}\|$.

Nous appellerons *angle des vecteurs non nuls* $\mathbf{x}$ *et* $\mathbf{y}$ l'angle de deux flèches d'origine commune représentant l'une $\mathbf{x}$ et l'autre $\mathbf{y}$.

2.1.4 Produit scalaire

On appelle *produit scalaire* des vecteurs non nuls x et y le nombre

$$\| \mathbf{x} \| \| \mathbf{y} \| \cos\theta, \tag{2.1}$$

où θ est l'angle de x et y. Si x ou y sont nuls, le produit scalaire est nul par convention.

Nous noterons le produit scalaire de x et y par $(\mathbf{x} \mid \mathbf{y})$. D'autres symboles couramment utilisés sont $(\mathbf{x}, \mathbf{y})$, $\langle \mathbf{x}, \mathbf{y} \rangle$, $\mathbf{x} \cdot \mathbf{y}$.

On remarquera que $(\mathbf{x} \mid \mathbf{x}) = \| \mathbf{x} \|^2$.

2.1.5 Orthogonalité

On dit que des vecteurs x et y sont *orthogonaux* s'ils sont non nuls et leur angle est $\frac{\pi}{2}$, ou si l'un d'entre eux est nul.

En vertu de la définition 2.1.4, x et y sont donc orthogonaux si et seulement si $(\mathbf{x} \mid \mathbf{y}) = 0$.

2.1.6 Projection orthogonale

La *projection orthogonale* d'un vecteur non nul x sur la droite vectorielle engendrée par un vecteur non nul v est le vecteur

$$\| \mathbf{x} \| \cos\theta \left(\frac{1}{\| \mathbf{v} \|} \mathbf{v}\right),$$

où θ est l'angle de x et v. Nous la noterons $\mathrm{proj}_{\mathbf{v}}\mathbf{x}$ et l'appellerons aussi projection orthogonale de x sur v (fig. 2.1).

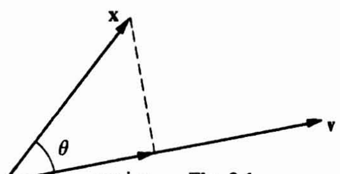

Fig. 2.1

A l'aide du produit scalaire, le vecteur $\mathrm{proj}_{\mathbf{v}}\mathbf{x}$ peut être exprimé autrement: il suffit de substituer $\frac{(\mathbf{x} \mid \mathbf{v})}{\| \mathbf{x} \| \| \mathbf{v} \|}$ à $\cos\theta$ pour obtenir

$$\mathrm{proj}_{\mathbf{v}}\mathbf{x} = \frac{(\mathbf{x} \mid \mathbf{v})}{\| \mathbf{v} \|^2} \mathbf{v} = \frac{(\mathbf{x} \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})} \mathbf{v}. \tag{2.2}$$

Cette expression vaut également dans le cas où x est nul, à condition d'admettre que la projection orthogonale du vecteur nul est le vecteur nul. La norme de $\mathrm{proj}_{\mathbf{v}}\mathbf{x}$ s'écrit

$$\| \mathrm{proj}_{\mathbf{v}}\mathbf{x} \| = \frac{|(\mathbf{x} \mid \mathbf{v})|}{\| \mathbf{v} \|^2} \| \mathbf{v} \| = \frac{|(\mathbf{x} \mid \mathbf{v})|}{\| \mathbf{v} \|}. \tag{2.3}$$

Si v est unitaire, (2.2) et (2.3) se simplifient et deviennent

$$\mathrm{proj}_{\mathbf{v}}\mathbf{x} = (\mathbf{x} \mid \mathbf{v})\mathbf{v}, \quad \| \mathrm{proj}_{\mathbf{v}}\mathbf{x} \| = |(\mathbf{x} \mid \mathbf{v})|. \tag{2.4}$$

Par des considérations géométriques élémentaires, il est facile de se rendre compte que

$$\mathrm{proj}_{\mathbf{v}}(\mathbf{x} + \mathbf{y}) = \mathrm{proj}_{\mathbf{v}}\mathbf{x} + \mathrm{proj}_{\mathbf{v}}\mathbf{y}, \quad \mathrm{proj}_{\mathbf{v}}\alpha\mathbf{x} = \alpha\,\mathrm{proj}_{\mathbf{v}}\mathbf{x}. \tag{2.5}$$

2.1.7 Propriétés

Le produit scalaire jouit de trois propriétés qui constitueront le point de départ d'une formulation abstraite de cette opération. Les voici:

(a) $(\mathbf{x} \mid \mathbf{y}) = (\mathbf{y} \mid \mathbf{x})$.
(b) $(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{z}) = \alpha(\mathbf{x} \mid \mathbf{z}) + \beta(\mathbf{y} \mid \mathbf{z})$.
(c) $(\mathbf{x} \mid \mathbf{x}) > 0$ si $\mathbf{x} \neq \mathbf{0}$.

La seule qui demande une vérification est la deuxième. Si z est nul, les trois produits scalaires sont nuls et l'égalité est vraie. Si z n'est pas nul,

$$\mathrm{proj}_{\mathbf{z}}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\,\mathrm{proj}_{\mathbf{z}}\mathbf{x} + \beta\,\mathrm{proj}_{\mathbf{z}}\mathbf{y},$$

d'après (2.5), ce qui entraîne, grâce à (2.2),

$$\frac{(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{z})}{(\mathbf{z} \mid \mathbf{z})}\mathbf{z} = \alpha\frac{(\mathbf{x} \mid \mathbf{z})}{(\mathbf{z} \mid \mathbf{z})}\mathbf{z} + \beta\frac{(\mathbf{y} \mid \mathbf{z})}{(\mathbf{z} \mid \mathbf{z})}\mathbf{z},$$

d'où la propriété (b) s'ensuit.

2.1.8 Bases orthonormales

Une base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ de V est dite orthonormale si les vecteurs $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ sont orthogonaux deux à deux et unitaires.

2.1.9 Décomposition suivant une base orthonormale

Par le raisonnement géométrique, on voit facilement que chaque vecteur est la somme de ses projections orthogonales sur les vecteurs d'une base orthonormale, autrement dit, si $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est une base orthonormale,

$$\mathbf{x} = (\mathbf{x} \mid \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{x} \mid \mathbf{e}_2)\mathbf{e}_2 + (\mathbf{x} \mid \mathbf{e}_3)\mathbf{e}_3. \tag{2.6}$$

Cette décomposition s'obtient également par (b) de 2.1.7. En effet, x_1, x_2, x_3 étant les composantes de **x**,

$$(\mathbf{x} \mid \mathbf{e}_1) = (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \mid \mathbf{e}_1) = x_1(\mathbf{e}_1 \mid \mathbf{e}_1) + x_2(\mathbf{e}_2 \mid \mathbf{e}_1) + x_3(\mathbf{e}_3 \mid \mathbf{e}_1) = x_1,$$

puisque $(\mathbf{e}_1 \mid \mathbf{e}_1) = 1$ et $(\mathbf{e}_2 \mid \mathbf{e}_1) = (\mathbf{e}_3 \mid \mathbf{e}_1) = 0$; de même,

$$(\mathbf{x} \mid \mathbf{e}_2) = x_2, \quad (\mathbf{x} \mid \mathbf{e}_3) = x_3,$$

d'où la décomposition.

2.1.10 Produit scalaire en fonction des composantes

Soit x_1, x_2, x_3 et y_1, y_2, y_3 les composantes respectives des vecteurs **x** et **y** dans une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Grâce à (a) et (b) de 2.1.7, le produit scalaire

$$(\mathbf{x} \mid \mathbf{y}) = (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \mid y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3)$$

se développe en une somme de neuf termes de la forme $x_iy_j(\mathbf{e}_i \mid \mathbf{e}_j)$; or, par l'orthonormalité de la base, $(\mathbf{e}_i \mid \mathbf{e}_j)$ est nul si $i \neq j$ et vaut 1 si $i = j$, ce qui entraîne

$$(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3.$$

2.2 ESPACES VECTORIELS EUCLIDIENS

2.2.1 Introduction

Le produit scalaire dans V a été défini au moyen des notions de norme et d'angle. Il jouit des propriétés (a), (b) et (c) (mises en évidence dans 2.1.7) qui en résument les caractères. Pour étendre la notion de produit scalaire aux espaces vectoriels abstraits, nous suivrons le procédé inverse; plus exactement, nous admettrons les trois propriétés comme données de départ, en déduirons les notions de norme, d'orthogonalité et d'angle et aboutirons à un certain nombre de résultats applicables à des situations les plus variées, notamment aux espaces vectoriels fonctionnels. La géométrie aura ainsi laissé la place à l'algèbre, en demeurant toutefois l'inspiratrice des idées et des méthodes utilisées.

2.2.2 Produit scalaire

Soit E un espace vectoriel. On appelle *produit scalaire* dans E toute opération qui fait correspondre à chaque couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E un nombre, noté $(\mathbf{x} \mid \mathbf{y})$ et appelé *produit scalaire* de **x** et **y**, satisfaisant aux conditions suivantes:

(a) $(\mathbf{x} \mid \mathbf{y}) = (\mathbf{y} \mid \mathbf{x})$ (symétrie ou commutativité).
(b) $(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{z}) = \alpha(\mathbf{x} \mid \mathbf{z}) + \beta(\mathbf{y} \mid \mathbf{z})$ (linéarité).
(c) $(\mathbf{x} \mid \mathbf{x}) > 0$ si $\mathbf{x} \neq \mathbf{0}$ (positivité).

2.2.3 Espaces vectoriels euclidiens

On appelle *espace vectoriel euclidien* tout espace vectoriel muni d'un produit scalaire.

Il est clair que tout sous-espace vectoriel d'un espace vectoriel euclidien E devient lui-même un espace vectoriel euclidien s'il est muni du produit scalaire hérité de E.

2.2.4 Remarques

(1) La condition (b) exprime la *linéarité à gauche* du produit scalaire (cf. 6.2.2). La *linéarité à droite* découle de l'union de (a) et (b):

$$(\mathbf{x} \mid \beta\mathbf{y} + \gamma\mathbf{z}) = \beta(\mathbf{x} \mid \mathbf{y}) + \gamma(\mathbf{x} \mid \mathbf{z}).$$

(2) En posant $\alpha = \beta = 0$ dans (b), ou $\beta = \gamma = 0$ ci-dessus, nous voyons que le produit scalaire $(\mathbf{x} \mid \mathbf{y})$ est nul si **x** ou **y** sont nuls.

(3) La linéarité à gauche et à droite du produit scalaire s'étend par récurrence aux combinaisons linéaires de plus de deux termes.

2.2.5 Norme d'un vecteur

Soit E un espace vectoriel euclidien. On appelle *norme* d'un vecteur **x** de E, et l'on note $\| \mathbf{x} \|$, le nombre $\sqrt{(\mathbf{x} \mid \mathbf{x})}$.

En vertu de la condition (c), la norme $\| \mathbf{x} \|$ est positive si **x** est non nul; d'après (2) de 2.2.4, elle est nulle si **x** est nul.

On remarquera que $\sqrt{(\alpha\mathbf{x} \mid \alpha\mathbf{x})} = \sqrt{\alpha^2(\mathbf{x} \mid \mathbf{x})} = |\alpha| \sqrt{(\mathbf{x} \mid \mathbf{x})}$, ce qui montre que

$$\| \alpha\mathbf{x} \| = |\alpha| \| \mathbf{x} \|. \tag{2.7}$$

On dit qu'un vecteur est *unitaire* si sa norme est 1. D'après (2.7), si **x** est non nul, $\dfrac{1}{\| \mathbf{x} \|}\mathbf{x}$ est un vecteur unitaire.

2.2.6 Exemples

(1) Vu les propriétés 2.1.7, l'opération définie dans 2.1.4 est un produit scalaire dans V conforme à la définition 2.2.2.

• (2) L'opération $(\cdot|\cdot)$ définie par la formule

$$\left(\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}\right) = a_1b_1 + a_2b_2 + \dots + a_nb_n \tag{2.8}$$

est un produit scalaire dans $\mathbb{R}^n$ que l'on appelle *produit scalaire canonique.*

Par la suite, sauf mention explicite du contraire, $\mathbb{R}^n$ sera considéré comme muni de ce produit scalaire.

Deux autres exemples de produits scalaires dans $\mathbb{R}^3$ sont définis par les formules

$$\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = 2a_1b_1 + 3a_2b_2 + 4a_3b_3,$$

$$\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = 7a_1b_1 + 2a_1b_2 + 2a_2b_1 + 6a_2b_2 + 2a_2b_3 + 2a_3b_2 + 5a_3b_3. \tag{2.9}$$

La preuve que l'opération définie par (2.9) satisfait à la condition (c) n'est pas immédiate. Nous laissons toutefois au lecteur la tâche de l'effectuer, car nous reviendrons sur cette question dans la section 9.2.

(3) Le champ d'application privilégié de la théorie abstraite du produit scalaire est constitué par les espaces vectoriels fonctionnels (cf. 1.3.4 et 1.3.5). On appelle *produit scalaire canonique* dans $C_{[a,b]}$ l'opération $(\cdot|\cdot)$ définie par la formule

$$(\mathbf{f}\,|\,\mathbf{g}) = \int_a^b \mathbf{f}(t)\mathbf{g}(t)\mathrm{d}t. \tag{2.10}$$

Cette opération définit bien un produit scalaire, car les conditions (a) et (b) de la définition 2.2.2 sont manifestement satisfaites et, en outre, l'intégrale

$$\int_a^b \mathbf{f}^2(t)\mathrm{d}t$$

est positive si la fonction continue $\mathbf{f}$ n'est pas identiquement nulle.

Par la suite, sauf mention explicite du contraire, $C_{[a,b]}$ sera considéré comme muni du produit scalaire canonique.

(4) Un autre exemple de produit scalaire dans $C_{[a,b]}$ est fourni par la formule

$$(\mathbf{f}\,|\,\mathbf{g}) = \int_a^b \mathbf{f}(t)\mathbf{g}(t)\mathbf{p}(t)\mathrm{d}t, \tag{2.11}$$

où $\mathbf{p}$ est une fonction continue à valeurs positives.

2.3 ORTHOGONALITÉ

2.3.1 Introduction

Il a été remarqué dans 2.1.5 que deux vecteurs de V sont orthogonaux si et seulement si leur produit scalaire est nul. Cette équivalence tiendra lieu de définition de la notion abstraite d'orthogonalité.

Dans la suite de ce chapitre, E désignera un espace vectoriel euclidien. Comme déjà convenu, sauf indication contraire, les vecteurs seront les éléments de E.

2.3.2 Orthogonalité

On dit que les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont *orthogonaux,* ou que $\mathbf{x}$ est orthogonal à $\mathbf{y}$, si $(\mathbf{x}\,|\,\mathbf{y}) = 0$. On dit qu'une famille finie ou infinie de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, \dots)$ est *orthogonale,* ou que les vecteurs $\mathbf{x}_1, \mathbf{x}_2, \dots$ sont *orthogonaux deux à deux,* si $(\mathbf{x}_i\,|\,\mathbf{x}_j) = 0$ pour tout couple (i, j) tel que $i \neq j$. Si, de plus, tous les vecteurs $\mathbf{x}_i$ sont unitaires, on dit que la famille est *orthonormale.*

Comme déjà noté dans la remarque 2.2.4, le vecteur nul est orthogonal à tout vecteur. En fait, vu la condition (c) de la définition 2.2.2, il est le seul vecteur qui possède cette propriété.

2.3.3 Exemples

(1) La base canonique de $\mathbb{R}^n$ est manifestement orthonormale.

(2) Considérons les fonctions $\mathbf{c}_0$, $\mathbf{c}_k$ et $\mathbf{s}_k$ $(k > 0)$ de $C_{[-\pi,\pi]}$ définies par

$$\mathbf{c}_0(t) \equiv \frac{1}{\sqrt{2\pi}}, \quad \mathbf{c}_k(t) \equiv \frac{1}{\sqrt{\pi}}\cos(kt), \quad \mathbf{s}_k(t) \equiv \frac{1}{\sqrt{\pi}}\sin(kt). \tag{2.12}$$

La famille infinie $(\mathbf{c}_0, \mathbf{c}_1, \mathbf{s}_1, \mathbf{c}_2, \mathbf{s}_2, \dots)$ est appelée *système trigonométrique.* Cette famille est orthonormale, car

$$\int_{-\pi}^{\pi} \cos(kt)\cos(lt)\mathrm{d}t = \tfrac{1}{2}\int_{-\pi}^{\pi} (\cos((k+l)t) + \cos((k-l)t))\mathrm{d}t = \begin{cases} 0 & \text{si } k \neq l, \\ \pi & \text{si } k = l \neq 0, \\ 2\pi & \text{si } k = l = 0; \end{cases}$$

$$\int_{-\pi}^{\pi} \sin(kt)\sin(lt)\mathrm{d}t = \tfrac{1}{2}\int_{-\pi}^{\pi} (\cos((k-l)t) - \cos((k+l)t))\mathrm{d}t = \begin{cases} 0 & \text{si } k \neq l, \\ \pi & \text{si } k = l \neq 0; \end{cases}$$

$$\int_{-\pi}^{\pi} \sin(kt)\cos(lt)\mathrm{d}t = \tfrac{1}{2}\int_{-\pi}^{\pi} (\sin((k+l)t) + \sin((k-l)t))\mathrm{d}t = 0.$$

2.3.4 Orthogonalité à un sous-espace vectoriel

On dit qu'un vecteur $\mathbf{x}$ est orthogonal à un sous-espace vectoriel S de E s'il est orthogonal à tout vecteur de S.

Si $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ est une famille génératrice de S, pour qu'un vecteur $\mathbf{x}$ soit orthogonal à S, il suffit qu'il soit orthogonal à tous les vecteurs $\mathbf{v}_i$. En effet, tout vecteur $\mathbf{y}$ de S s'écrit sous la forme $\mathbf{y} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k$, donc, par la linéarité à droite du produit scalaire,

$$\begin{aligned}(\mathbf{x} \mid \mathbf{y}) &= (\mathbf{x} \mid \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k) \\ &= \alpha_1(\mathbf{x} \mid \mathbf{v}_1) + \alpha_2(\mathbf{x} \mid \mathbf{v}_2) + ... + \alpha_k(\mathbf{x} \mid \mathbf{v}_k) = 0,\end{aligned}$$

ce qui démontre l'assertion.

2.3.5 Sous-espaces vectoriels orthogonaux

On dit que des sous-espaces vectoriels S et T de E sont orthogonaux, ou que S est orthogonal à T, si chaque vecteur de S est orthogonal à T (ou, ce qui revient au même, chaque vecteur de T est orthogonal à S).

2.3.6 Proposition. Orthogonalité et indépendance linéaire

Toute famille orthogonale finie de vecteurs non nuls est libre. En particulier, toute famille orthonormale finie est libre.

DÉMONSTRATION

Soit $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ une famille orthogonale de vecteurs non nuls. La relation $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = 0$ entraîne, pour $i = 1, 2, ..., k$, grâce à la linéarité à gauche du produit scalaire,

$$\begin{aligned}0 = (\mathbf{0} \mid \mathbf{x}_i) &= (\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k \mid \mathbf{x}_i) \\ &= \alpha_1(\mathbf{x}_1 \mid \mathbf{x}_i) + \alpha_2(\mathbf{x}_2 \mid \mathbf{x}_i) + ... + \alpha_k(\mathbf{x}_k \mid \mathbf{x}_i) = \alpha_i(\mathbf{x}_i \mid \mathbf{x}_i),\end{aligned}$$

d'où nous concluons que $\alpha_i = 0$, puisque $(\mathbf{x}_i \mid \mathbf{x}_i) \neq 0$.

2.3.7 Théorème de Pythagore

La relation

$$(\mathbf{x} + \mathbf{y} \mid \mathbf{x} + \mathbf{y}) = (\mathbf{x} \mid \mathbf{x}) + (\mathbf{y} \mid \mathbf{y}) + 2(\mathbf{x} \mid \mathbf{y}) \tag{2.13}$$

résulte aisément de la linéarité à gauche et à droite du produit scalaire. Cette relation s'écrit aussi sous la forme

$$\| \mathbf{x} + \mathbf{y} \|^2 = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 + 2(\mathbf{x} \mid \mathbf{y}) \tag{2.14}$$

et devient, lorsque $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux, le *théorème de Pythagore*:

$$\| \mathbf{x} + \mathbf{y} \|^2 = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2. \tag{2.15}$$

L'extension de (2.15) au cas d'une famille orthogonale $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ de plus de deux termes se fait par récurrence:

$$\| \mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_k \|^2 = \| \mathbf{x}_1 \|^2 + \| \mathbf{x}_2 \|^2 + ... + \| \mathbf{x}_k \|^2. \tag{2.16}$$

2.3.8 Projection orthogonale sur une droite vectorielle

On appelle *projection orthogonale d'un vecteur* $\mathbf{x}$ sur la droite vectorielle engendrée par un vecteur non nul $\mathbf{v}$, ou simplement projection orthogonale de $\mathbf{x}$ sur $\mathbf{v}$, le vecteur

$$\operatorname{proj}_{\mathbf{v}}\mathbf{x} = \frac{(\mathbf{x} \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})}\mathbf{v}. \tag{2.17}$$

On remarquera que cette projection est l'unique vecteur $\alpha\mathbf{v}$ de la droite en question tel que $\mathbf{x} - \alpha\mathbf{v}$ est orthogonal à $\mathbf{v}$. En effet, l'unique solution de l'équation $(\mathbf{x} - \alpha\mathbf{v} \mid \mathbf{v}) = 0$ est

$$\alpha = \frac{(\mathbf{x} \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})}.$$

2.3.9 Orthogonalisation

Le procédé que nous allons présenter est connu sous le nom de *procédé d'orthogonalisation de Gram-Schmidt.* A partir d'une famille libre $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$, il montre par quelles opérations sur les vecteurs $\mathbf{x}_i$ on peut construire une famille orthogonale $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ engendrant le même sous-espace vectoriel que $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$.

On pose successivement

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{x}_1, \\ \mathbf{v}_2 &= \mathbf{x}_2 - \frac{(\mathbf{x}_2 \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1, \\ \mathbf{v}_3 &= \mathbf{x}_3 - \frac{(\mathbf{x}_3 \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 - \frac{(\mathbf{x}_3 \mid \mathbf{v}_2)}{(\mathbf{v}_2 \mid \mathbf{v}_2)}\mathbf{v}_2, \\ &\ldots\ldots \\ \mathbf{v}_k &= \mathbf{x}_k - \frac{(\mathbf{x}_k \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 - \frac{(\mathbf{x}_k \mid \mathbf{v}_2)}{(\mathbf{v}_2 \mid \mathbf{v}_2)}\mathbf{v}_2 - ... - \frac{(\mathbf{x}_k \mid \mathbf{v}_{k-1})}{(\mathbf{v}_{k-1} \mid \mathbf{v}_{k-1})}\mathbf{v}_{k-1}.\end{aligned} \tag{2.18}$$

En d'autres termes, $\mathbf{v}_i$ $(i > 1)$ est la différence de $\mathbf{x}_i$ et la somme des projections orthogonales de $\mathbf{x}_i$ sur les vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_{i-1}$.

Nous allons montrer, par récurrence, que les vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ sont orthogonaux deux à deux et non nuls. A cet effet, supposons que les vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i$ $(i < k)$ le soient. Alors, pour $j = 1, 2, \ldots, i$,

$$\begin{aligned}(\mathbf{v}_{i+1} \mid \mathbf{v}_j) &= (\mathbf{x}_{i+1} - \sum_{l=1}^{i} \frac{(\mathbf{x}_{i+1} \mid \mathbf{v}_l)}{(\mathbf{v}_l \mid \mathbf{v}_l)} \mathbf{v}_l \mid \mathbf{v}_j) \\ &= (\mathbf{x}_{i+1} \mid \mathbf{v}_j) - \sum_{l=1}^{i} \frac{(\mathbf{x}_{i+1} \mid \mathbf{v}_l)}{(\mathbf{v}_l \mid \mathbf{v}_l)} (\mathbf{v}_l \mid \mathbf{v}_j) = (\mathbf{x}_{i+1} \mid \mathbf{v}_j) - \frac{(\mathbf{x}_{i+1} \mid \mathbf{v}_j)}{(\mathbf{v}_j \mid \mathbf{v}_j)} (\mathbf{v}_j \mid \mathbf{v}_j) = 0,\end{aligned}$$

ce qui démontre que $\mathbf{v}_{i+1}$ est orthogonal à $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i$. En outre, $\mathbf{v}_{i+1}$ n'est pas nul, sinon $\mathbf{x}_{i+1}$ serait combinaison linéaire des vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i$ et donc des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i$, en contradiction avec l'hypothèse d'indépendance linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. Les vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i+1}$ sont donc orthogonaux deux à deux et non nuls, ce qui établit notre assertion.

On notera que, pour chaque i, les familles $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i)$ et $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i)$ engendrent le même sous-espace vectoriel de E.

2.3.10 Exemples

(1) Par le procédé de Gram-Schmidt (2.18), nous allons orthogonaliser le triplet de $\mathbb{R}^4$

$$\left(\begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}\right).$$

En effectuant les calculs, nous obtenons

$$\mathbf{v}_1 = \begin{pmatrix}1\\1\\1\\0\end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix}1\\-1\\1\\1\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\1\\1\\0\end{pmatrix} = \frac{1}{3}\begin{pmatrix}2\\-4\\2\\3\end{pmatrix} \quad \text{(le facteur } \frac{1}{3} \text{ peut être omis)},$$

$$\mathbf{v}_3 = \begin{pmatrix}1\\1\\-1\\-1\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\1\\1\\0\end{pmatrix} - \frac{-7}{33}\begin{pmatrix}2\\-4\\2\\3\end{pmatrix} = -\frac{2}{11}\begin{pmatrix}-6\\1\\5\\2\end{pmatrix} \quad \text{(le facteur } -\frac{2}{11} \text{ peut être omis)}.$$

(2) Soit $(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2)$ la famille de vecteurs de $C_{[-1,1]}$ définie par $\mathbf{p}_i(t) \equiv t^i$ (cf. (1.8) et exemple (5) de 1.6.7). Le procédé d'orthogonalisation de Gram-Schmidt (2.18) appliqué à cette famille entraîne:

$$\begin{aligned}\mathbf{v}_0(t) &\equiv 1,\\ \mathbf{v}_1(t) &\equiv t - \tfrac{1}{2}\int_{-1}^{1} s\,ds \equiv t,\\ \mathbf{v}_2(t) &\equiv t^2 - \tfrac{1}{2}\int_{-1}^{1} s^2 ds - \tfrac{3}{2}\int_{-1}^{1} s^3 ds \cdot t \equiv t^2 - \tfrac{1}{3}.\end{aligned}$$

Les polynômes $\mathbf{v}_0$, $\mathbf{v}_1$ et $\mathbf{v}_2$ ainsi obtenus sont les trois premiers termes d'une famille orthogonale infinie connue sous le nom de *système de Legendre*.

2.4 INÉGALITÉS, ANGLES

2.4.1 Inégalité de Cauchy-Schwarz

On appelle *inégalité de Cauchy-Schwarz* l'inégalité, valable pour tout choix des vecteurs $\mathbf{x}$ et $\mathbf{y}$,

$$|(\mathbf{x} \mid \mathbf{y})| \leqslant \|\mathbf{x}\| \|\mathbf{y}\|. \tag{2.19}$$

Dans le cas particulier où $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs de V et le produit scalaire est défini par (2.1), cette inégalité est évidente, puisque $|\cos\theta| \leqslant 1$. Dans le cas général, elle nécessite une preuve. Si $\mathbf{y}$ est nul, les deux membres de (2.19) sont nuls et l'inégalité est en fait une égalité. Si $\mathbf{y}$ n'est pas nul, posons $\mathbf{e} = \frac{1}{\|\mathbf{y}\|}\mathbf{y}$ et écrivons

$$\mathbf{x} = (\mathbf{x} \mid \mathbf{e})\mathbf{e} + \mathbf{x} - (\mathbf{x} \mid \mathbf{e})\mathbf{e}. \tag{2.20}$$

Comme $(\mathbf{x} \mid \mathbf{e})\mathbf{e}$ est la projection orthogonale de $\mathbf{x}$ sur $\mathbf{e}$, $\mathbf{x} - (\mathbf{x} \mid \mathbf{e})\mathbf{e}$ est orthogonal à $\mathbf{e}$, donc à $(\mathbf{x} \mid \mathbf{e})\mathbf{e}$, ce qui nous permet d'appliquer (2.15) au second membre de (2.20) et d'obtenir

$$\|\mathbf{x}\|^2 = \|(\mathbf{x} \mid \mathbf{e})\mathbf{e}\|^2 + \|\mathbf{x} - (\mathbf{x} \mid \mathbf{e})\mathbf{e}\|^2 \geqslant (\mathbf{x} \mid \mathbf{e})^2\|\mathbf{e}\|^2 = (\mathbf{x} \mid \mathbf{e})^2. \tag{2.21}$$

Il ne reste alors plus qu'à multiplier les racines carrées des deux membres extrêmes de (2.21) par $\|\mathbf{y}\|$ et à substituer $\mathbf{y}$ à $\|\mathbf{y}\|\mathbf{e}$ pour établir (2.19).

On remarquera que l'inégalité de Cauchy-Schwarz est une égalité si et seulement si $\mathbf{x}$ et $\mathbf{y}$ sont linéairement dépendants. En effet, l'inégalité dans (2.21) est une égalité si et seulement si $\|\mathbf{x} - (\mathbf{x} \mid \mathbf{e})\mathbf{e}\|$ est nul, autrement dit, $\mathbf{x}$ est multiple de $\mathbf{e}$ ou, ce qui revient au même, de $\mathbf{y}$.

2.4.2 Exemples d'application

(1) Lorsque E est $\mathbb{R}^n$, l'inégalité de Cauchy-Schwarz s'écrit

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leqslant \sqrt{\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2}.$$

Dans le cas particulier où $b_1 = b_2 = \ldots = b_n = 1$, elle devient

$$\left(\sum_{i=1}^{n} a_i\right)^2 \leqslant n\sum_{i=1}^{n} a_i^2,$$

ou encore

$$\left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)^2 \leqslant \frac{1}{n}\sum_{i=1}^{n} a_i^2,$$

ce qui montre que le carré de la moyenne arithmétique est inférieur ou égal à la moyenne arithmétique des carrés.

(2) Lorsque E est $C_{[a,b]}$, l'inégalité de Cauchy-Schwarz s'écrit

$$\left|\int_a^b \mathbf{f}(t)\mathbf{g}(t)\mathrm{d}t\right| \leqslant \sqrt{\int_a^b \mathbf{f}^2(t)\mathrm{d}t\int_a^b \mathbf{g}^2(t)\mathrm{d}t}.$$

En prenant $\mathbf{g}(t) \equiv 1$ et $|\mathbf{f}|$ à la place de $\mathbf{f}$, nous en déduisons l'inégalité

$$\left(\int_a^b |\mathbf{f}(t)|\mathrm{d}t\right)^2 \leqslant (b-a)\int_a^b \mathbf{f}^2(t)\mathrm{d}t.$$

Par exemple,

$$\left(\int_0^\pi \sqrt{\sin t}\,\mathrm{d}t\right)^2 \leqslant \pi\int_0^\pi \sin t\,\mathrm{d}t = 2\pi,$$

ou encore

$$\int_0^\pi \sqrt{\sin t}\,\mathrm{d}t \leqslant \sqrt{2\pi}.$$

2.4.3 Inégalité triangulaire

En majorant $2(\mathbf{x}\,|\,\mathbf{y})$ par $2\|\mathbf{x}\|\|\mathbf{y}\|$ (inégalité de Cauchy-Schwarz) dans (2.14), nous voyons que

$$\|\mathbf{x}+\mathbf{y}\|^2 \leqslant \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,$$

ce qui entraîne *l'inégalité triangulaire*:

$$\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|. \tag{2.22}$$

En l'appliquant une fois aux vecteurs $\mathbf{x}$ et $\mathbf{y}-\mathbf{x}$ et une autre fois aux vecteurs $\mathbf{x}-\mathbf{y}$ et $\mathbf{y}$, nous déduisons la variante

$$|\,\|\mathbf{x}\| - \|\mathbf{y}\|\,| \leqslant \|\mathbf{x}-\mathbf{y}\|.$$

2.4.4 Angles. Théorème du cosinus

Supposons que les vecteurs $\mathbf{x}$ et $\mathbf{y}$ soient non nuls. En vertu de l'inégalité de Cauchy-Schwarz,

$$-1 \leqslant \frac{(\mathbf{x}\,|\,\mathbf{y})}{\|\mathbf{x}\|\|\mathbf{y}\|} \leqslant 1.$$

Il existe donc un et un seul nombre θ de l'intervalle $[0, \pi]$ tel que

$$\cos\theta = \frac{(\mathbf{x}\,|\,\mathbf{y})}{\|\mathbf{x}\|\|\mathbf{y}\|}. \tag{2.23}$$

Ce nombre est appelé *angle des vecteurs* $\mathbf{x}$ *et* $\mathbf{y}$. On notera les cas particuliers suivants:

$$\mathbf{x} \text{ et } \mathbf{y} \text{ orthogonaux: } \frac{(\mathbf{x}\,|\,\mathbf{y})}{\|\mathbf{x}\|\|\mathbf{y}\|} = 0, \text{ ce qui équivaut à } \theta = \frac{\pi}{2};$$

$$\mathbf{x} = \beta\mathbf{y}\text{: } \frac{(\mathbf{x}\,|\,\mathbf{y})}{\|\mathbf{x}\|\|\mathbf{y}\|} = \frac{\beta\|\mathbf{y}\|^2}{|\beta|\|\mathbf{y}\|^2} = \begin{cases} 1 \text{ si } \beta > 0, \text{ ce qui équivaut à } \theta = 0; \\ -1 \text{ si } \beta < 0, \text{ ce qui équivaut à } \theta = \pi. \end{cases}$$

Si $\mathbf{x}$ et $\mathbf{y}$ sont non nuls, la relation (2.14), ou celle qui en dérive par substitution de $-\mathbf{y}$ à $\mathbf{y}$, s'écrit, à l'aide de (2.23), sous la forme équivalente

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \pm 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta, \tag{2.24}$$

où θ est l'angle de $\mathbf{x}$ et $\mathbf{y}$. On appelle (2.24) *théorème du cosinus*.

2.5 ESPACES VECTORIELS EUCLIDIENS DE DIMENSION FINIE

2.5.1 Introduction

Cette section a pour objet l'étude de quelques questions liées à l'existence d'une base orthonormale.

Nous supposerons que l'espace vectoriel euclidien E est de dimension finie non nulle n.

2.5.2 Proposition. Prolongement d'une famille orthonormale

Soit $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k)$ *une famille de vecteurs orthonormale. Si* k *est inférieur à la dimension* n *de* E, *cette famille se prolonge en une base orthonormale de* E.

DÉMONSTRATION

D'après le théorème 1.7.3, la famille $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$ se prolonge en une base $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{x}_{k+1}, ..., \mathbf{x}_n)$ de E. Appliqué à cette base, le procédé d'orthogonalisation de Gram-Schmidt (2.18) nous fournit une base orthogonale $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{v}_{k+1}, ..., \mathbf{v}_n)$ de E. Posons

$$\mathbf{e}_{k+1} = \frac{1}{\| \mathbf{v}_{k+1} \|}\mathbf{v}_{k+1}, ..., \mathbf{e}_n = \frac{1}{\| \mathbf{v}_n \|}\mathbf{v}_n.$$

La famille $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est alors une base orthonormale de E.

2.5.3 Corollaire. Existence d'une base orthonormale

Tout espace vectoriel euclidien de dimension finie non nulle admet une base orthonormale.

DÉMONSTRATION

Un tel espace comprend au moins un vecteur non nul $\mathbf{e}_1$, que nous pouvons supposer unitaire. Il suffit alors d'appliquer la proposition 2.5.2 à la famille $(\mathbf{e}_1)$.

2.5.4 Composantes d'un vecteur dans une base orthonormale

Soit $x_1, x_2, ..., x_n$ les composantes d'un vecteur $\mathbf{x}$ dans une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. Par la linéarité à gauche du produit scalaire, pour $i = 1, 2, ..., n$,

$$\begin{aligned}(\mathbf{x} \mid \mathbf{e}_i) &= (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n \mid \mathbf{e}_i)\\ &= x_1(\mathbf{e}_1 \mid \mathbf{e}_i) + x_2(\mathbf{e}_2 \mid \mathbf{e}_i) + ... + x_n(\mathbf{e}_n \mid \mathbf{e}_i)\\ &= x_i(\mathbf{e}_i \mid \mathbf{e}_i) = x_i,\end{aligned}$$

ce qui montre que $x_i\mathbf{e}_i$ est la projection orthogonale de $\mathbf{x}$ sur $\mathbf{e}_i$. La décomposition de $\mathbf{x}$ suivant la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ s'écrit donc

$$\mathbf{x} = (\mathbf{x} \mid \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{x} \mid \mathbf{e}_2)\mathbf{e}_2 + ... + (\mathbf{x} \mid \mathbf{e}_n)\mathbf{e}_n. \tag{2.25}$$

Cela généralise la décomposition (2.6).

2.5.5 Expression du produit scalaire en fonction des composantes

Soit $x_1, x_2, ..., x_n$ et $y_1, y_2, ..., y_n$ les composantes respectives de $\mathbf{x}$ et de $\mathbf{y}$ dans une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. Par la linéarité à gauche et à droite du produit scalaire,

$$(\mathbf{x} \mid \mathbf{y}) = \left(\sum_{i=1}^{n} x_i\mathbf{e}_i \,\Big|\, \sum_{j=1}^{n} y_j\mathbf{e}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j(\mathbf{e}_i \mid \mathbf{e}_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j, \tag{2.26}$$

où $a_{ij} = (\mathbf{e}_i \mid \mathbf{e}_j)$. En particulier,

$$(\mathbf{x} \mid \mathbf{x}) = \| \mathbf{x} \|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j. \tag{2.27}$$

Nous reviendrons sur les expressions ainsi obtenues dans le chapitre 9. Pour l'instant, limitons-nous à observer que si la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormale, (2.26) et (2.27) se simplifient et se réduisent aux expressions qui définissent le produit scalaire canonique dans $\mathbb{R}^n$ et le carré de la norme correspondante:

$$(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 + ... + x_ny_n, \tag{2.28}$$

$$(\mathbf{x} \mid \mathbf{x}) = \| \mathbf{x} \|^2 = x_1^2 + x_2^2 + ... + x_n^2. \tag{2.29}$$

2.5.6 Isomorphie de E et $\mathbb{R}^n$

D'après (2.28), lorsque la base de E est orthonormale, l'isomorphisme (1.11) entre E et $\mathbb{R}^n$ conserve le produit scalaire, ce qui signifie que le produit scalaire de $\mathbf{x}$ et $\mathbf{y}$ dans E est égal au produit scalaire dans $\mathbb{R}^n$ des vecteurs-colonnes correspondants (x_i) et (y_i).

2.5.7 Produit scalaire défini par une base

Tout espace vectoriel de dimension finie non nulle n peut être rendu euclidien. Il suffit de choisir une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de cet espace et de définir le produit scalaire par la formule (2.28). Pour ce produit scalaire, la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormale.

Il est évident que tout espace vectoriel réduit au vecteur nul peut également être considéré comme euclidien.

2.6 PROJECTION ORTHOGONALE ET MEILLEURE APPROXIMATION

2.6.1 Introduction

Dans cette section, nous définirons la notion de projection orthogonale d'un vecteur sur un sous-espace vectoriel et utiliserons cette notion pour calculer la meilleure approximation d'un vecteur par des vecteurs d'un sous-espace vectoriel de dimension finie.

L'espace vectoriel euclidien E sera de nouveau de dimension finie ou infinie.

2.6.2 Complémentaire orthogonal

Soit S un sous-espace vectoriel de E. Désignons par $S^\perp$ l'ensemble des vecteurs orthogonaux à S. Cet ensemble n'est pas vide, car il comprend au moins le vecteur nul. En outre, par la linéarité du produit scalaire, en même temps que $\mathbf{x}$ et $\mathbf{y}$ il comprend toutes les combinaisons linéaires $\alpha\mathbf{x} + \beta\mathbf{y}$. D'après la proposition 1.4.6, $S^\perp$ est donc un sous-espace vectoriel de E. Ce sous-espace est orthogonal à S, selon la définition 2.3.5.

D'après (c) de la définition 2.2.2, $S \cap S^\perp = \{\mathbf{0}\}$. Cependant, E n'est pas, en général, somme directe de S et $S^\perp$ (cf. exercice 2.9.13). Il l'est, toutefois, s'il existe un sous-espace vectoriel T de E, orthogonal à S et tel que E soit somme directe de S et T, car alors T est nécessairement $S^\perp$. En effet, grâce à l'existence d'un tel T, tout vecteur $\mathbf{t}'$ de $S^\perp$ s'écrit sous la forme $\mathbf{t}' = \mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T; cela entraîne que $\mathbf{t}' - \mathbf{t}$ est un vecteur de S orthogonal à S, donc que $\mathbf{t}' - \mathbf{t} = \mathbf{0}$ et, par conséquent, que $\mathbf{t}'$ ($= \mathbf{t}$) appartient à T, ce qui montre que $S^\perp$ est contenu dans T. D'autre part, par définition de $S^\perp$, T est contenu dans $S^\perp$, donc $T = S^\perp$.

Il s'avère ainsi que les conditions suivantes sont équivalentes:

(a) Il existe un sous-espace vectoriel T de E, orthogonal à S et tel que E soit somme directe de S et T.
(b) E est somme directe de S et $S^\perp$.
(c) Tout vecteur $\mathbf{x}$ de E s'écrit (nécessairement de manière unique) sous la forme
$$\mathbf{x} = \mathbf{s} + \mathbf{t}, \tag{2.30}$$
où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur orthogonal à S.
(d) Pour tout vecteur $\mathbf{x}$ de E, il existe un vecteur $\mathbf{s}$ de S (nécessairement unique) tel que $\mathbf{x} - \mathbf{s}$ est orthogonal à S.

On dit que S *admet un complémentaire orthogonal dans* E, ou que $S^\perp$ *est le complémentaire orthogonal de* S *dans* E, si l'une des conditions équivalentes (a)–(d) est satisfaite.

D'après ce que nous venons d'établir, si $S^\perp$ est le complémentaire orthogonal de S dans E, alors S est le complémentaire orthogonal de $S^\perp$ dans E et
$$(S^\perp)^\perp = S. \tag{2.31}$$

2.6.3 Projection orthogonale

Soit S un sous-espace vectoriel de E et $\mathbf{x}$ un vecteur de E. Si S admet un complémentaire orthogonal dans E, on appelle le vecteur $\mathbf{s}$ de la décomposition (2.30) *projection orthogonale de* $\mathbf{x}$ *sur* S. Nous désignerons cette projection par $\mathrm{proj}_S\mathbf{x}$.

2.6.4 Proposition. Existence du complémentaire orthogonal d'un sous-espace vectoriel de dimension finie

Tout sous-espace vectoriel de E de dimension finie admet un complémentaire orthogonal dans E.

DÉMONSTRATION

Soit S un sous-espace vectoriel de E de dimension finie. Soit en outre $\mathbf{x}$ un vecteur quelconque de E. Nous allons montrer qu'il existe un vecteur $\mathbf{s}$ de S tel que $\mathbf{x} - \mathbf{s}$ est orthogonal à S. Si S se réduit à $\{\mathbf{0}\}$, $\mathbf{s}$ est le vecteur nul. Si S est de dimension non nulle k, choisissons une base orthogonale quelconque $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ de S et posons
$$\mathbf{s} = \mathrm{proj}_{\mathbf{v}_1}\mathbf{x} + ... + \mathrm{proj}_{\mathbf{v}_k}\mathbf{x} = \frac{(\mathbf{x} \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 + ... + \frac{(\mathbf{x} \mid \mathbf{v}_k)}{(\mathbf{v}_k \mid \mathbf{v}_k)}\mathbf{v}_k.$$
Compte tenu de 2.3.8, nous voyons que $\mathbf{x} - \mathbf{s}$ est orthogonal à tous les vecteurs $\mathbf{v}_i$. D'après 2.3.4, $\mathbf{x} - \mathbf{s}$ est donc orthogonal à S.

2.6.5 Calcul de la projection orthogonale

Lorsqu'on doit calculer la projection orthogonale d'un vecteur $\mathbf{x}$ de E sur un sous-espace vectoriel S de E défini par une de ses bases $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$, il faut avoir en vue les deux cas suivants:

(a) La base $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ est orthogonale. Dans ce cas, d'après ce qui précède,
$$\mathrm{proj}_S\mathbf{x} = \frac{(\mathbf{x} \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 + \frac{(\mathbf{x} \mid \mathbf{v}_2)}{(\mathbf{v}_2 \mid \mathbf{v}_2)}\mathbf{v}_2 + ... + \frac{(\mathbf{x} \mid \mathbf{v}_k)}{(\mathbf{v}_k \mid \mathbf{v}_k)}\mathbf{v}_k. \tag{2.32}$$

(b) La base $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ n'est pas orthogonale. Dans ce cas, soit on l'orthogonalise par le procédé de Gram-Schmidt et on applique (2.32) à la base orthogonale $(\mathbf{v}_1', \mathbf{v}_2', ..., \mathbf{v}_k')$ obtenue par ce procédé, soit on détermine les coefficients de la combinaison linéaire $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k$ de manière que le vecteur $\mathbf{x} - (\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k)$ soit orthogonal à $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$; cette condition se traduit par les k équations
$$\alpha_1(\mathbf{v}_1 \mid \mathbf{v}_i) + \alpha_2(\mathbf{v}_2 \mid \mathbf{v}_i) + ... + \alpha_k(\mathbf{v}_k \mid \mathbf{v}_i) = (\mathbf{x} \mid \mathbf{v}_i), \quad i = 1, 2, ..., k,$$
d'où l'on tire les valeurs de $\alpha_1, \alpha_2, ..., \alpha_k$ qui produisent la combinaison linéaire égale à $\mathrm{proj}_S\mathbf{x}$.

2.6.6 Vecteur normal à un hyperplan vectoriel

Soit S un hyperplan vectoriel de E (supposé de dimension finie non nulle). D'après la proposition 2.6.4, S admet un complémentaire orthogonal dans E. D'après (1.12), ce complémentaire est une droite vectorielle. On appelle *vecteur normal* à S tout vecteur non nul de cette droite vectorielle.

2.6.7 Meilleure approximation

Soit S un sous-espace vectoriel de E de dimension finie. On appelle *meilleure approximation* d'un vecteur $\mathbf{x}$ par des vecteurs de S l'unique vecteur de S qui minimise la norme $\| \mathbf{x} - \mathbf{y} \|$, lorsque $\mathbf{y}$ parcourt S. Dans le cas particulier où E est l'espace vectoriel géométrique et S une droite ou un plan vectoriel de cet espace, on voit facilement que la meilleure approximation de $\mathbf{x}$ est la projection orthogonale de $\mathbf{x}$ sur S. Nous allons montrer qu'il en va de même dans le cas général, c'est-à-dire que

$$\| \mathbf{x} - \operatorname{proj}_S \mathbf{x} \| < \| \mathbf{x} - \mathbf{y} \|, \tag{2.33}$$

pour tout vecteur $\mathbf{y}$ de S distinct de $\operatorname{proj}_S \mathbf{x}$. Considérons un tel vecteur $\mathbf{y}$ et écrivons

$$\mathbf{x} - \mathbf{y} = (\mathbf{x} - \operatorname{proj}_S \mathbf{x}) + (\operatorname{proj}_S \mathbf{x} - \mathbf{y}).$$

Comme $\mathbf{x} - \operatorname{proj}_S \mathbf{x}$ est orthogonal à tout vecteur de S, donc à $\operatorname{proj}_S \mathbf{x} - \mathbf{y}$, la relation (2.15) entraîne

$$\| \mathbf{x} - \mathbf{y} \|^2 = \| \mathbf{x} - \operatorname{proj}_S \mathbf{x} \|^2 + \| \operatorname{proj}_S \mathbf{x} - \mathbf{y} \|^2 > \| \mathbf{x} - \operatorname{proj}_S \mathbf{x} \|^2,$$

ce qui établit (2.33).

2.6.8 Exemples d'application

La meilleure approximation est utilisée notamment dans des problèmes de prédiction et d'interpolation. Les deux exemples que nous présentons concernent le système de Legendre et le système trigonométrique.

(1) La meilleure approximation de la fonction $\mathbf{f}(t) \equiv |t|$ de $C_{[-1,1]}$ par des polynômes de degré inférieur ou égal à 2 est, d'après 2.6.7 et (2.32), le polynôme

$$\mathbf{p} = \frac{(\mathbf{f} \mid \mathbf{v}_0)}{(\mathbf{v}_0 \mid \mathbf{v}_0)}\mathbf{v}_0 + \frac{(\mathbf{f} \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 + \frac{(\mathbf{f} \mid \mathbf{v}_2)}{(\mathbf{v}_2 \mid \mathbf{v}_2)}\mathbf{v}_2,$$

où $\mathbf{v}_0$, $\mathbf{v}_1$ et $\mathbf{v}_2$ sont les polynômes orthogonaux obtenus dans l'exemple (2) de 2.3.10. Le calcul des produits scalaires donne:

$$(\mathbf{f} \mid \mathbf{v}_0) = \int_{-1}^{1} |t| \mathrm{d}t = 1, \quad (\mathbf{v}_0 \mid \mathbf{v}_0) = \int_{-1}^{1} 1^2 \mathrm{d}t = 2;$$

$$(\mathbf{f} \mid \mathbf{v}_1) = \int_{-1}^{1} |t| t \mathrm{d}t = 0 \quad \text{(le calcul de } (\mathbf{v}_1 \mid \mathbf{v}_1) \text{ est donc inutile)};$$

$$(\mathbf{f} \mid \mathbf{v}_2) = \int_{-1}^{1} |t|(t^2 - \tfrac{1}{3}) \mathrm{d}t = \tfrac{1}{6}, \quad (\mathbf{v}_2 \mid \mathbf{v}_2) = \int_{-1}^{1} (t^2 - \tfrac{1}{3})^2 \mathrm{d}t = \tfrac{8}{45}.$$

Il en résulte donc que

$$\mathbf{p}(t) \equiv \tfrac{1}{2} + \tfrac{1}{6} \cdot \tfrac{45}{8}(t^2 - \tfrac{1}{3}) \equiv \tfrac{3}{16}(5t^2 + 1) \quad \text{(fig. 2.2)}.$$

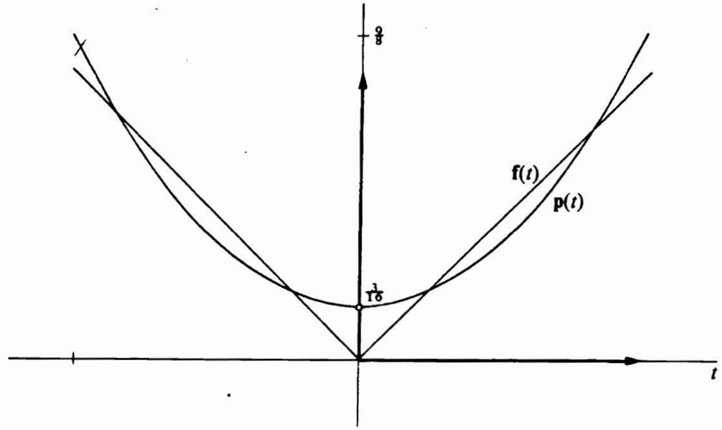

Fig. 2.2

(2) La meilleure approximation de la fonction $\mathbf{f}(t) \equiv t$ de $C_{[-\pi,\pi]}$ par des *polynômes trigonométriques* d'ordre inférieur ou égal à k, c'est-à-dire des combinaisons linéaires des $2k + 1$ premiers termes du système trigonométrique défini dans 2.3.3, est, d'après 2.6.7 et (2.32), le polynôme trigonométrique

$$\mathbf{f}_k = (\mathbf{f} \mid \mathbf{c}_0)\mathbf{c}_0 + (\mathbf{f} \mid \mathbf{c}_1)\mathbf{c}_1 + (\mathbf{f} \mid \mathbf{s}_1)\mathbf{s}_1 + \ldots + (\mathbf{f} \mid \mathbf{c}_k)\mathbf{c}_k + (\mathbf{f} \mid \mathbf{s}_k)\mathbf{s}_k.$$

Les produits scalaires $(\mathbf{f} \mid \mathbf{c}_j)$ $(j \geqslant 0)$ et $(\mathbf{f} \mid \mathbf{s}_j)$ $(j \geqslant 1)$ sont appelés *coefficients de Fourier* de la fonction $\mathbf{f}$. Comme $\mathbf{f} \cdot \mathbf{c}_j$ est une fonction impaire,

$$(\mathbf{f} \mid \mathbf{c}_j) = \frac{1}{\sqrt{(2)\pi}} \int_{-\pi}^{\pi} t \cos(jt) \mathrm{d}t = 0.$$

D'autre part,

$$(\mathbf{f} \mid \mathbf{s}_j) = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} t \sin(jt) \mathrm{d}t = \frac{2\sqrt{\pi}}{j}(-1)^{j+1},$$

donc le polynôme trigonométrique $\mathbf{f}_k$ s'écrit sous la forme

$$\mathbf{f}_k(t) \equiv \sum_{j=1}^{k} \frac{2}{j}(-1)^{j+1} \sin(jt).$$

L'étude de la convergence de $\mathbf{f}_k(t)$, lorsque k tend vers l'infini, dépasse le cadre de ce livre et sera donc omise. Signalons néanmoins qu'elle a lieu et que la limite est $t = \mathbf{f}(t)$ si $-\pi < t < \pi$ et 0 si $t = \pm\pi$ (car $\sin(j\pi) = 0$). D'une manière générale, on peut démontrer que $\mathbf{f}_k$ *converge en norme* vers $\mathbf{f}$, c'est-à-dire que

$$\lim_{k \to \infty} \| \mathbf{f} - \mathbf{f}_k \| = 0. \tag{2.34}$$

2.7 PRODUIT VECTORIEL ET PRODUIT MIXTE

2.7.1 Introduction

Le produit vectoriel de deux vecteurs est une opération propre à la dimension 3. Pour l'introduire, il faut préalablement orienter l'espace destiné à le recevoir. L'orientation étant définie au moyen de la notion de déterminant, nous commencerons par une brève introduction à l'étude de cette notion. Cette étude sera reprise dans le chapitre 5.

Dans cette section, nous supposerons que l'espace vectoriel euclidien E est de dimension 3.

2.7.2 Déterminants d'ordre 2 et 3

On appelle *déterminant* des vecteurs-colonnes de $\mathbb{R}^2$

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

et on note

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \tag{2.35}$$

le nombre

$$a_1b_2 - a_2b_1. \tag{2.36}$$

On appelle *déterminant* des vecteurs-colonnes de $\mathbb{R}^3$

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix},$$

et on note

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \tag{2.37}$$

le nombre

$$a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \tag{2.38}$$

$$= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1.$$

La fonction qui associe à tout couple de vecteurs-colonnes de $\mathbb{R}^2$ (à tout triplet de vecteurs-colonnes de $\mathbb{R}^3$) son déterminant est appelée déterminant d'ordre 2 (d'ordre 3).

2.7.3 Propriétés des déterminants

Les propriétés les plus importantes des déterminants seront établies dans la section 5.1. En anticipant sur cette section, nous énonçons ici les deux propriétés qui nous servent à étudier le produit vectoriel et le produit mixte.

(a) Le déterminant est multiplié par -1 si l'un de ses vecteurs-colonnes est remplacé par son opposé ou si deux de ses vecteurs-colonnes sont échangés.
(b) Le déterminant est non nul si et seulement si ses vecteurs-colonnes sont linéairement indépendants.

Ces deux propriétés sont évidentes dans le cas du déterminant d'ordre 2. La propriété (a) l'est également dans le cas du déterminant d'ordre 3. La vérification de la propriété (b) dans ce même cas est laissée comme exercice. S'il le désire, le lecteur pourra se reporter au paragraphe 5.2.2, où cette vérification est faite en toute généralité.

2.7.4 Déterminants de passage

On appelle *déterminant de passage* d'une base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ à une base $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$ le déterminant

$$\begin{vmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{vmatrix}, \tag{2.39}$$

où p_{1j}, p_{2j}, p_{3j} sont les composantes du vecteur $\mathbf{y}_j$ dans la base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$.

2.7.5 Orientation de E

On définit l'orientation de l'espace vectoriel E par le choix d'une de ses bases. Ce choix étant fait, on dit que E est *orienté* ou *doté d'une orientation.* Soit $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ la base définissant l'orientation de E. On dit alors qu'une base $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$ est *directe* ou *positivement orientée* si le déterminant de passage (2.39) est positif et *indirecte* ou *négativement orientée* si ce même déterminant est négatif.

2.7.6 Remarque

Nous verrons dans 6.6.10 que le déterminant de passage entre bases de même orientation est positif, tandis qu'entre bases d'orientations différentes ce déterminant est négatif. Par les déterminants de passage, les bases de E peuvent donc être rangées en deux classes définissant chacune l'une des deux orientations possibles de E.

2.7.7 Effet d'une transposition et d'une permutation cyclique

Les deux assertions suivantes découlent directement de (a) de 2.7.3.

(1) L'orientation d'une base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ change si l'un des vecteurs $\mathbf{x}_i$ est remplacé par son opposé, ou si deux vecteurs $\mathbf{x}_i$ et $\mathbf{x}_j$ sont échangés.

(2) L'orientation d'une base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ reste inchangée si les vecteurs $\mathbf{x}_i$ sont permutés cycliquement, c'est-à-dire si $\mathbf{x}_1$ prend la place de $\mathbf{x}_2$, $\mathbf{x}_2$ celle de $\mathbf{x}_3$ et $\mathbf{x}_3$ celle de $\mathbf{x}_1$ (ou, ce qui revient au même, si $\mathbf{x}_1$ et $\mathbf{x}_3$ et ensuite $\mathbf{x}_2$ et $\mathbf{x}_1$ sont échangés).

A l'aide de (1) et (2), un examen des six dispositions possibles de $\mathbf{x}_1$, $\mathbf{x}_2$ et $\mathbf{x}_3$ nous permet de conclure que les trois triplets

$$(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3), \quad (\mathbf{x}_3, \mathbf{x}_1, \mathbf{x}_2), \quad (\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_1)$$

sont de même orientation, tandis que les trois autres

$$(\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_3), \quad (\mathbf{x}_3, \mathbf{x}_2, \mathbf{x}_1), \quad (\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_2)$$

sont d'orientation opposée à celle des trois précédents.

2.7.8 Choix d'une base

Dans la suite de cette section, $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est une base orthonormale définissant l'orientation de E.

2.7.9 Produit vectoriel

Soit x_1, x_2, x_3 et y_1, y_2, y_3 les composantes respectives des vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. On appelle *produit vectoriel* de $\mathbf{x}$ et $\mathbf{y}$, et on note $\mathbf{x} \times \mathbf{y}$, le vecteur

$$(x_2y_3 - x_3y_2)\mathbf{e}_1 + (x_3y_1 - x_1y_3)\mathbf{e}_2 + (x_1y_2 - x_2y_1)\mathbf{e}_3. \tag{2.40}$$

Cette définition dépend du choix de la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, mais nous verrons dans 2.7.13 qu'elle n'en dépend pas si ce choix est fait parmi les bases orthonormales directes.

La i-ième composante de $\mathbf{x} \times \mathbf{y}$ est le déterminant des deux colonnes

$$\begin{array}{cc} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{array} \tag{2.41}$$

privées de leur i-ième terme, le deuxième déterminant étant cependant pris avec le signe $-$:

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\mathbf{e}_3. \tag{2.42}$$

Cas particuliers:

$$\mathbf{e}_1 \times \mathbf{e}_1 = \mathbf{e}_2 \times \mathbf{e}_2 = \mathbf{e}_3 \times \mathbf{e}_3 = \mathbf{0},$$
$$\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, \quad \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \quad \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2.$$

2.7.10 Propriétés du produit vectoriel

Le produit vectoriel jouit des propriétés suivantes:

(a) $\mathbf{x} \times \mathbf{y} = -(\mathbf{y} \times \mathbf{x})$ (antisymétrie).
(b) $(\alpha\mathbf{x} + \beta\mathbf{y}) \times \mathbf{z} = \alpha(\mathbf{x} \times \mathbf{z}) + \beta(\mathbf{y} \times \mathbf{z})$ (linéarité).
(c) $\mathbf{x} \times \mathbf{y} \neq \mathbf{0}$ si et seulement si $\mathbf{x}$ et $\mathbf{y}$ sont linéairement indépendants.

Les deux premières découlent directement de la définition. Nous établirons la troisième en montrant que $\mathbf{x} \times \mathbf{y}$ est nul si et seulement si $\mathbf{x}$ et $\mathbf{y}$ sont linéairement dépendants. Sous cette condition, l'une des deux colonnes dans (2.41) est multiple de l'autre, donc les trois déterminants dans (2.42) sont nuls, ce qui entraîne que le produit vectoriel $\mathbf{x} \times \mathbf{y}$ est nul. Réciproquement, supposons que $\mathbf{x} \times \mathbf{y}$ soit nul. Alors les trois déterminants dans (2.42) sont nuls, c'est-à-dire

$$x_2y_3 = x_3y_2, \quad x_3y_1 = x_1y_3, \quad x_1y_2 = x_2y_1. \tag{2.43}$$

Si tous les y_i sont nuls, les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont linéairement dépendants. Si un des y_i, par exemple y_3, est non nul, il existe α tel que $x_3 = \alpha y_3$, ce qui entraîne, par substitution de αy_3 à x_3 dans (2.43),

$$x_2y_3 = \alpha y_3y_2, \quad \alpha y_3y_1 = x_1y_3$$

et donc

$$x_2 = \alpha y_2, \quad x_1 = \alpha y_1.$$

La première colonne dans (2.41) est donc le produit de α par la deuxième, ce qui implique la dépendance linéaire de $\mathbf{x}$ et $\mathbf{y}$.

La propriété (b) exprime la *linéarité à gauche* du produit vectoriel (cf. 6.2.2). Cette propriété, jointe à l'antisymétrie, entraîne la *linéarité à droite*:

$$\mathbf{x} \times (\beta\mathbf{y} + \gamma\mathbf{z}) = \beta(\mathbf{x} \times \mathbf{y}) + \gamma(\mathbf{x} \times \mathbf{z}). \tag{2.44}$$

On remarquera, en particulier, que

$$\alpha(\mathbf{x} \times \mathbf{y}) = (\alpha\mathbf{x}) \times \mathbf{y} = \mathbf{x} \times (\alpha\mathbf{y}), \tag{2.45}$$

ce qui nous permet d'omettre les parenthèses et d'écrire $\alpha\mathbf{x} \times \mathbf{y}$ sans danger d'ambiguïté.

2.7.11 Propriétés métriques du produit vectoriel

Les vecteurs $\mathbf{x}$ et $\mathbf{y}$ étant supposés non nuls, le produit vectoriel de $\mathbf{x}$ et $\mathbf{y}$ est

(a) orthogonal à $\mathbf{x}$ et à $\mathbf{y}$,
(b) de norme $\|\mathbf{x}\|\,\|\mathbf{y}\|\sin\theta$, où θ est l'angle de $\mathbf{x}$ et $\mathbf{y}$.

En effet, d'après (2.42), (2.28), 2.7.2 et (b) de 2.7.3,

$$(\mathbf{x} \mid \mathbf{x} \times \mathbf{y}) = x_1 \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} - x_2 \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} + x_3 \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & x_1 & y_1 \\ x_2 & x_2 & y_2 \\ x_3 & x_3 & y_3 \end{vmatrix} = 0,$$

ce qui montre que $\mathbf{x} \times \mathbf{y}$ est orthogonal à $\mathbf{x}$ et donc aussi à $\mathbf{y}$, puisque $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$. D'autre part, d'après (2.40), (2.29), (2.28) et (2.23),

$$\begin{aligned} \| \mathbf{x} \times \mathbf{y} \|^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\ &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\ &= \| \mathbf{x} \|^2 \| \mathbf{y} \|^2 \sin^2\theta, \end{aligned}$$

ce qui démontre (b).

On remarquera que dans le cas où E est l'espace vectoriel géométrique V^3, la norme $\| \mathbf{x} \| \| \mathbf{y} \| \sin\theta$ du produit vectoriel représente l'aire du parallélogramme construit sur des représentants de $\mathbf{x}$ et $\mathbf{y}$ d'origine commune (fig. 2.3).

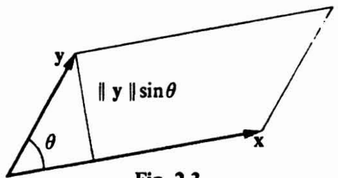

Fig. 2.3

2.7.12 Orientation

Si $\mathbf{x}$ et $\mathbf{y}$ sont linéairement indépendants, le triplet $(\mathbf{x} \times \mathbf{y}, \mathbf{x}, \mathbf{y})$ et donc aussi le triplet $(\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y})$ sont directs.

En effet, z_1, z_2, z_3 étant les composantes de $\mathbf{x} \times \mathbf{y}$ (dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$), le déterminant de passage de $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ à $(\mathbf{x} \times \mathbf{y}, \mathbf{x}, \mathbf{y})$ s'écrit

$$\begin{vmatrix} z_1 & x_1 & y_1 \\ z_2 & x_2 & y_2 \\ z_3 & x_3 & y_3 \end{vmatrix} = z_1^2 + z_2^2 + z_3^2. \tag{2.46}$$

Ce déterminant est donc positif, puisqu'au moins un des z_i n'est pas nul, d'après (c) de 2.7.10.

2.7.13 Indépendance du produit vectoriel du choix de la base orthonormale directe

L'expression du produit vectoriel (2.40) est invariante par changement de base orthonormale directe. Soit en effet $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$ une telle base. Selon les propriétés métriques 2.7.11, $\mathbf{e}'_1 \times \mathbf{e}'_2$ est soit $\mathbf{e}'_3$, soit $-\mathbf{e}'_3$. D'après 2.7.12, le triplet $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_1 \times \mathbf{e}'_2)$ est direct, donc $\mathbf{e}'_1 \times \mathbf{e}'_2 = \mathbf{e}'_3$, vu l'assertion (1) de 2.7.7. Par le même

raisonnement, compte tenu de l'assertion (2) de 2.7.7, nous voyons que $\mathbf{e}'_2 \times \mathbf{e}'_3 = \mathbf{e}'_1$ et $\mathbf{e}'_3 \times \mathbf{e}'_1 = \mathbf{e}'_2$. Soit maintenant x'_1, x'_2, x'_3 et y'_1, y'_2, y'_3 les composantes respectives de $\mathbf{x}$ et $\mathbf{y}$ dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$. Grâce à (a) et (b) de 2.7.10, nous pouvons écrire

$$\begin{aligned} \mathbf{x} \times \mathbf{y} &= \left(\sum_{i=1}^{3} x'_i \mathbf{e}'_i\right) \times \left(\sum_{j=1}^{3} y'_j \mathbf{e}'_j\right) = \sum_{i=1}^{3}\sum_{j=1}^{3} x'_i y'_j \mathbf{e}'_i \times \mathbf{e}'_j \\ &= (x'_2y'_3 - x'_3y'_2)\mathbf{e}'_1 + (x'_3y'_1 - x'_1y'_3)\mathbf{e}'_2 + (x'_1y'_2 - x'_2y'_1)\mathbf{e}'_3, \end{aligned}$$

ce qui montre l'invariance de (2.40).

2.7.14 Produit mixte

On appelle *produit mixte* des vecteurs $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ le double produit $(\mathbf{x} \mid \mathbf{y} \times \mathbf{z})$.

D'après (2.23), si $\mathbf{x}$ et $\mathbf{y} \times \mathbf{z}$ sont non nuls, le produit mixte de $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ s'écrit aussi sous la forme

$$\| \mathbf{x} \| \| \mathbf{y} \times \mathbf{z} \| \cos\theta, \tag{2.47}$$

où θ est l'angle de $\mathbf{x}$ et $\mathbf{y} \times \mathbf{z}$.

On remarquera que dans le cas où E est l'espace vectoriel géométrique V^3, le produit mixte, écrit sous la forme (2.47), symbolise le volume (orienté) du parallélépipède construit sur des représentants de $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ d'origine commune (fig. 2.4).

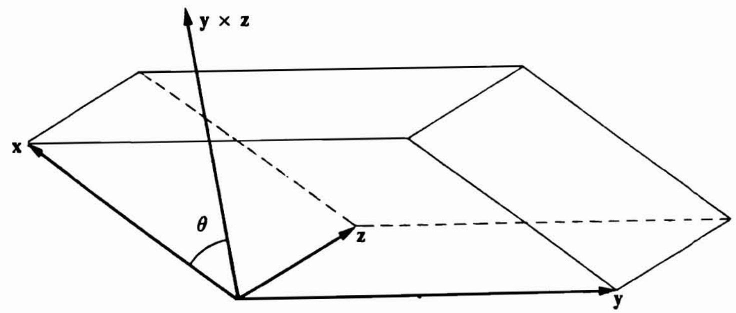

Fig. 2.4

2.7.15 Produit mixte en fonction des composantes

Soit x_1, x_2, x_3, y_1, y_2, y_3 et z_1, z_2, z_3 les composantes respectives de $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ dans une base orthonormale directe. D'après (2.42), (2.28) et 2.7.2,

$$(\mathbf{x} \mid \mathbf{y} \times \mathbf{z}) = x_1 \begin{vmatrix} y_2 & z_2 \\ y_3 & z_3 \end{vmatrix} - x_2 \begin{vmatrix} y_1 & z_1 \\ y_3 & z_3 \end{vmatrix} + x_3 \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}, \tag{2.48}$$

donc l'expression du produit mixte en fonction des composantes est le déterminant des vecteurs-colonnes des composantes. Il s'ensuit, grâce à la commutativité du produit scalaire et à (a) de 2.7.3, que

$$(\mathbf{x} \times \mathbf{y} \mid \mathbf{z}) = (\mathbf{z} \mid \mathbf{x} \times \mathbf{y}) = (\mathbf{x} \mid \mathbf{y} \times \mathbf{z}), \tag{2.49}$$

ce qui montre que l'ordre dans lequel les deux produits sont effectués est indifférent et justifie la notation suivante:

2.7.16 Notation

Désormais le produit mixte des vecteurs $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ sera noté $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$.

2.7.17 Propriétés du produit mixte

Le produit mixte jouit des propriétés suivantes:

(a) $[\mathbf{x}, \mathbf{y}, \mathbf{z}] = [\mathbf{z}, \mathbf{x}, \mathbf{y}] = [\mathbf{y}, \mathbf{z}, \mathbf{x}] = -[\mathbf{y}, \mathbf{x}, \mathbf{z}] = -[\mathbf{z}, \mathbf{y}, \mathbf{x}] = -[\mathbf{x}, \mathbf{z}, \mathbf{y}]$.
(b) $[\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}, \mathbf{v}] = \alpha[\mathbf{x}, \mathbf{z}, \mathbf{v}] + \beta[\mathbf{y}, \mathbf{z}, \mathbf{v}]$.
(c) $[\mathbf{x}, \mathbf{y}, \mathbf{z}] \neq 0$ si et seulement si $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ sont linéairement indépendants.

La première et la troisième de ces propriétés découlent de la relation (2.48) jointe respectivement à (a) et à (b) de 2.7.3. La deuxième résulte de la définition 2.7.14 jointe à (b) de 2.2.2.

En particulier, la propriété (a) montre que le produit mixte est invariant par permutation cyclique.

Par (2.48), la propriété (b) résulte aussi immédiatement de (2.38). Elle exprime la *linéarité à gauche* du produit mixte (cf. 6.2.2). La *linéarité au centre* et *à droite* suit de l'union de (a) et (b).

2.7.18 Quelques identités vectorielles

Voici trois identités vectorielles d'utilité pratique:

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = (\mathbf{x} \mid \mathbf{z})\mathbf{y} - (\mathbf{x} \mid \mathbf{y})\mathbf{z}, \tag{2.50}$$
$$(\mathbf{x} \times \mathbf{y} \mid \mathbf{z} \times \mathbf{v}) = (\mathbf{x} \mid \mathbf{z})(\mathbf{y} \mid \mathbf{v}) - (\mathbf{x} \mid \mathbf{v})(\mathbf{y} \mid \mathbf{z}), \tag{2.51}$$
$$(\mathbf{x} \times \mathbf{y}) \times (\mathbf{z} \times \mathbf{v}) = [\mathbf{x}, \mathbf{y}, \mathbf{v}]\mathbf{z} - [\mathbf{x}, \mathbf{y}, \mathbf{z}]\mathbf{v}. \tag{2.52}$$

Le lecteur pourra vérifier aisément (2.50) au moyen des composantes dans une base orthonormale directe. Par (2.49), la première croix et la barre peuvent être échangées dans le premier membre de (2.51); grâce à (2.50), il s'ensuit que

$$\begin{aligned}(\mathbf{x} \times \mathbf{y} \mid \mathbf{z} \times \mathbf{v}) &= (\mathbf{x} \mid \mathbf{y} \times (\mathbf{z} \times \mathbf{v}))\\ &= (\mathbf{x} \mid (\mathbf{y} \mid \mathbf{v})\mathbf{z} - (\mathbf{y} \mid \mathbf{z})\mathbf{v}) = (\mathbf{x} \mid \mathbf{z})(\mathbf{y} \mid \mathbf{v}) - (\mathbf{x} \mid \mathbf{v})(\mathbf{y} \mid \mathbf{z}),\end{aligned}$$

ce qui établit (2.51). Pour obtenir (2.52), il suffit d'appliquer (2.50) au double produit vectoriel $\mathbf{u} \times (\mathbf{z} \times \mathbf{v})$, où $\mathbf{u} = \mathbf{x} \times \mathbf{y}$, car cela donne, compte tenu de (2.49),

$$(\mathbf{x} \times \mathbf{y}) \times (\mathbf{z} \times \mathbf{v}) = (\mathbf{x} \times \mathbf{y} \mid \mathbf{v})\mathbf{z} - (\mathbf{x} \times \mathbf{y} \mid \mathbf{z})\mathbf{v} = [\mathbf{x}, \mathbf{y}, \mathbf{v}]\mathbf{z} - [\mathbf{x}, \mathbf{y}, \mathbf{z}]\mathbf{v},$$

ce qui prouve (2.52).

L'identité (2.50) montre, en particulier, que le produit vectoriel n'est pas associatif, c'est-à-dire qu'en général

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) \neq (\mathbf{x} \times \mathbf{y}) \times \mathbf{z}.$$

2.8 ESPACES AFFINES EUCLIDIENS

2.8.1 Introduction

Le concept d'espace affine défini dans la section 1.9 met en relation un espace ponctuel avec la structure linéaire d'un espace vectoriel. Dans un espace affine, des notions comme la distance ou l'orthogonalité n'ont pas de sens. Pour leur en donner un, il faut munir l'espace vectoriel directeur d'un produit scalaire. Le but de cette section est précisément d'examiner les conséquences qu'implique l'introduction d'une telle opération.

2.8.2 Espaces affines euclidiens

Soit $\mathscr{E}$ un espace affine de direction E. On dit que $\mathscr{E}$ est un *espace affine euclidien* si E est un espace vectoriel euclidien.

En vertu de 1.10.5 et 2.2.3, tout sous-espace affine d'un espace affine euclidien est lui-même un espace affine euclidien.

2.8.3 Exemples

(1) L'espace $\mathscr{G}$ de la géométrie élémentaire est un espace affine euclidien, car l'espace directeur V qui lui est associé est un espace vectoriel euclidien.

(2) Dans l'exemple (2) de 1.9.7, il est montré que tout espace vectoriel E peut être considéré comme un espace affine d'espace directeur E lui-même. Cet espace affine est euclidien si E est euclidien.

(3) Tout espace affine de dimension finie peut être rendu euclidien par le procédé indiqué dans 2.5.7.

(4) L'ensemble des solutions du système (1.15) est un sous-espace affine euclidien de $\mathbb{R}^3$.

2.8.4 Distance et ses propriétés

Dans la suite de cette section, $\mathscr{E}$ désignera un espace affine euclidien. Le lecteur désireux de se familiariser de nouveau avec la notation est prié de se reporter à la section 1.9.

On appelle *distance des points P et Q* le nombre $\delta(P, Q)$ défini par

$$\delta(P, Q) = \| \overrightarrow{PQ} \|. \tag{2.53}$$

Des règles (1), (2) et (3) de 1.9.5 et des propriétés de la norme vues dans 2.2.5, il résulte que $\delta(P, Q) = \delta(Q, P)$, que $\delta(P, Q)$ est positif si les points P et Q sont distincts et nul s'ils sont confondus. De la condition (b) de la définition 1.9.2 et de (2.22), il résulte en outre que tout triplet de points (P, Q, R) vérifie l'inégalité triangulaire:

$$\delta(P, R) \leqslant \delta(P, Q) + \delta(Q, R). \tag{2.54}$$

Cette inégalité est une égalité si et seulement si P, Q et R sont alignés, c'est-à-dire appartiennent à une même droite.

2.8.5 Sous-espaces affines orthogonaux

On dit que deux sous-espaces affines de $\mathscr{E}$ sont orthogonaux, ou que l'un d'entre eux est orthogonal à l'autre, si leurs espaces directeurs sont orthogonaux.

2.8.6 Vecteur normal à un hyperplan

Lorsque $\mathscr{E}$ est de dimension finie non nulle, on appelle *vecteur normal* à un hyperplan $\mathscr{S}$ de $\mathscr{E}$ tout vecteur normal à la direction S de $\mathscr{S}$. Ainsi, une droite orthogonale à un hyperplan $\mathscr{S}$ a pour vecteur directeur un vecteur normal à $\mathscr{S}$.

2.8.7 Projection orthogonale d'un point

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et de dimension finie. Soit en outre P un point de $\mathscr{E}$. Nous allons démontrer qu'il existe un point R de $\mathscr{S}$ et un seul tel que $\overrightarrow{RP}$ soit orthogonal à S. Ce point R s'appelle *projection orthogonale de P sur* $\mathscr{S}$ et se note $\mathrm{proj}_{\mathscr{S}} P$.

Soit P_0 un point quelconque de $\mathscr{S}$. Posons $R = P_0 + \mathrm{proj}_S \overrightarrow{P_0P}$ et montrons que R est la projection cherchée (fig. 2.5). D'après (1.16), R est un point de $\mathscr{S}$; en outre, d'après (d) de 2.6.2, le vecteur

$$\overrightarrow{P_0P} - \mathrm{proj}_S \overrightarrow{P_0P} = \overrightarrow{P_0P} - \overrightarrow{P_0R} = \overrightarrow{RP}$$

est orthogonal à S. D'autre part, si Q est un point de $\mathscr{S}$ tel que $\overrightarrow{QP}$ est orthogonal à S, en écrivant

$$\overrightarrow{QR} = \overrightarrow{QP} - \overrightarrow{RP},$$

nous voyons que $\overrightarrow{QR}$ est un vecteur de S orthogonal à S, donc à lui-même. Par conséquent, $\overrightarrow{QR} = \mathbf{0}$, ce qui entraîne $Q = R$.

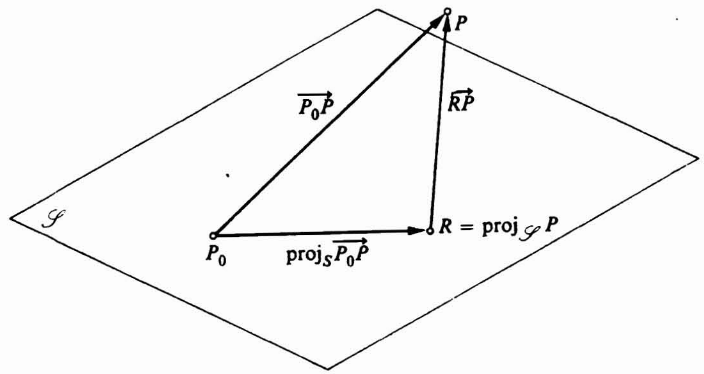

Fig. 2.5

2.8.8 Perpendiculaire par un point à un sous-espace affine

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de dimension finie. Soit en outre P un point de $\mathscr{E}$ n'appartenant pas à $\mathscr{S}$. On appelle *perpendiculaire* par P à $\mathscr{S}$ la droite déterminée par P et $\mathrm{proj}_{\mathscr{S}} P$.

2.8.9 Distance d'un point à un sous-espace affine

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de dimension finie. On appelle *distance d'un point P à* $\mathscr{S}$, et l'on note $\mathrm{dist}(P, \mathscr{S})$, le minimum de $\delta(P, Q)$, Q parcourant $\mathscr{S}$. Nous nous proposons de démontrer que

$$\mathrm{dist}(P, \mathscr{S}) = \delta(P, \mathrm{proj}_{\mathscr{S}} P). \tag{2.55}$$

Soit Q un point de $\mathscr{S}$ distinct de $\mathrm{proj}_{\mathscr{S}} P$. Comme $\overrightarrow{(\mathrm{proj}_{\mathscr{S}} P)Q}$ est orthogonal à $\overrightarrow{P(\mathrm{proj}_{\mathscr{S}} P)}$, par (2.15) nous voyons que

$$\begin{aligned} \delta^2(P, Q) &= \| \overrightarrow{PQ} \|^2 = \| \overrightarrow{P(\mathrm{proj}_{\mathscr{S}} P)} + \overrightarrow{(\mathrm{proj}_{\mathscr{S}} P)Q} \|^2 \\ &= \| \overrightarrow{P(\mathrm{proj}_{\mathscr{S}} P)} \|^2 + \| \overrightarrow{(\mathrm{proj}_{\mathscr{S}} P)Q} \|^2 > \| \overrightarrow{P(\mathrm{proj}_{\mathscr{S}} P)} \|^2 \\ &= \delta^2(P, \mathrm{proj}_{\mathscr{S}} P), \end{aligned}$$

ce qui prouve (2.55).

Lorsque P n'appartient pas à $\mathscr{S}$, la connaissance d'un vecteur directeur $\mathbf{n}$ de la perpendiculaire par P à $\mathscr{S}$ permet de calculer la distance de P à $\mathscr{S}$ au moyen de la formule

$$\operatorname{dist}(P, \mathscr{S}) = \| \operatorname{proj}_{\mathbf{n}} \overrightarrow{P_0P} \| = \frac{|(\overrightarrow{P_0P} \mid \mathbf{n})|}{\|\mathbf{n}\|}, \tag{2.56}$$

où P_0 est un point quelconque de $\mathscr{S}$. Pour justifier cette formule, il suffit d'observer que $\overrightarrow{P_0P} - \overrightarrow{(\operatorname{proj}_{\mathscr{S}} P)P} = \overrightarrow{P_0(\operatorname{proj}_{\mathscr{S}} P)}$ est un vecteur orthogonal à $\mathbf{n}$, ce qui nous permet de conclure que

$$\overrightarrow{(\operatorname{proj}_{\mathscr{S}} P)P} = \operatorname{proj}_{\mathbf{n}} \overrightarrow{P_0P} = \frac{(\overrightarrow{P_0P} \mid \mathbf{n})}{(\mathbf{n} \mid \mathbf{n})}\mathbf{n}.$$

Si $\mathscr{E}$ est de dimension finie non nulle et $\mathscr{S}$ est un hyperplan de $\mathscr{E}$, tout vecteur $\mathbf{n}$ normal à $\mathscr{S}$ est un vecteur directeur de la perpendiculaire par P à $\mathscr{S}$. Lorsque la dimension de $\mathscr{E}$ est 3, un tel vecteur est fourni par le produit vectoriel $\mathbf{v}_1 \times \mathbf{v}_2$, où $\mathbf{v}_1$, $\mathbf{v}_2$ sont des vecteurs directeurs de $\mathscr{S}$ (pour le cas général, voir l'exercice 5.5.13).

2.8.10 Perpendiculaire commune à deux droites gauches

Soit $\mathscr{D}$ et $\mathscr{D}'$ deux droites gauches de $\mathscr{E}$, c'est-à-dire deux droites non parallèles et d'intersection vide. Nous allons démontrer qu'il existe un unique point P de $\mathscr{D}$ et un unique point P' de $\mathscr{D}'$ tels que la droite déterminée par P et P' soit orthogonale à $\mathscr{D}$ et à $\mathscr{D}'$. Cette droite est appelée *perpendiculaire commune* à $\mathscr{D}$ et à $\mathscr{D}'$ (fig. 2.6).

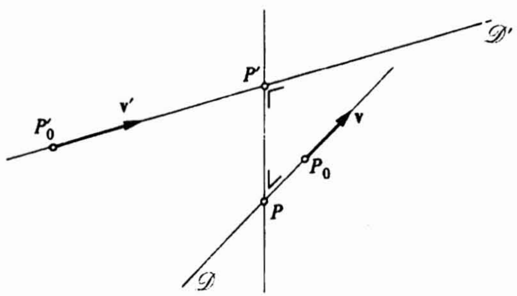

Fig. 2.6

Soit P_0 un point quelconque de $\mathscr{D}$ et P_0' un point quelconque de $\mathscr{D}'$. Soit en outre $\mathbf{v}$ un vecteur directeur de $\mathscr{D}$ et $\mathbf{v}'$ un vecteur directeur de $\mathscr{D}'$. Des représentations paramétriques de $\mathscr{D}$ et de $\mathscr{D}'$

$$P = P_0 + \alpha\mathbf{v} \quad \text{et} \quad P' = P_0' + \alpha'\mathbf{v}', \tag{2.57}$$

nous déduisons la relation

$$\overrightarrow{PP'} = \overrightarrow{P_0P_0'} + \alpha'\mathbf{v}' - \alpha\mathbf{v}.$$

Posons

$$\mathbf{u} = \mathbf{v} - \operatorname{proj}_{\mathbf{v}'}\mathbf{v} = \mathbf{v} - \frac{(\mathbf{v} \mid \mathbf{v}')}{(\mathbf{v}' \mid \mathbf{v}')}\mathbf{v}'. \tag{2.58}$$

On notera que $\mathbf{u}$ n'est pas nul, car $\mathscr{D}$ et $\mathscr{D}'$ ne sont pas parallèles. Supposons que la droite déterminée par P et P' soit orthogonale à $\mathscr{D}$ et à $\mathscr{D}'$. Alors le vecteur $\overrightarrow{PP'}$ est orthogonal à $\mathbf{v}$ et à $\mathbf{v}'$, donc à $\mathbf{u}$. En d'autres termes,

$$(\overrightarrow{P_0P_0'} \mid \mathbf{u}) + \alpha'(\mathbf{v}' \mid \mathbf{u}) - \alpha(\mathbf{v} \mid \mathbf{u}) = 0. \tag{2.59}$$

Mais

$$(\mathbf{v}' \mid \mathbf{u}) = (\mathbf{v}' \mid \mathbf{v}) - \frac{(\mathbf{v} \mid \mathbf{v}')}{(\mathbf{v}' \mid \mathbf{v}')}(\mathbf{v}' \mid \mathbf{v}') = 0$$

et

$$(\mathbf{v} \mid \mathbf{u}) = (\mathbf{v} \mid \mathbf{v}) - \frac{(\mathbf{v} \mid \mathbf{v}')}{(\mathbf{v}' \mid \mathbf{v}')}(\mathbf{v} \mid \mathbf{v}') = (\mathbf{u} \mid \mathbf{u}),$$

donc l'équation (2.59) se réduit à

$$(\overrightarrow{P_0P_0'} \mid \mathbf{u}) - \alpha(\mathbf{u} \mid \mathbf{u}) = 0.$$

En posant

$$\mathbf{u}' = \mathbf{v}' - \operatorname{proj}_{\mathbf{v}}\mathbf{v}' = \mathbf{v}' - \frac{(\mathbf{v}' \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})}\mathbf{v}, \tag{2.60}$$

nous obtenons, de manière analogue, l'équation

$$(\overrightarrow{P_0P_0'} \mid \mathbf{u}') + \alpha'(\mathbf{u}' \mid \mathbf{u}') = 0.$$

Nous en concluons que

$$\alpha = \frac{(\overrightarrow{P_0P_0'} \mid \mathbf{u})}{(\mathbf{u} \mid \mathbf{u})} \quad \text{et} \quad \alpha' = -\frac{(\overrightarrow{P_0P_0'} \mid \mathbf{u}')}{(\mathbf{u}' \mid \mathbf{u}')}, \tag{2.61}$$

où $\mathbf{u}$ et $\mathbf{u}'$ sont définis dans (2.58) et (2.60). Par les représentations (2.57), ces valeurs de α et de α' déterminent les points P et P' cherchés. En effet, le vecteur $\overrightarrow{PP'}$ est orthogonal à $\mathbf{u}$ et à $\mathbf{u}'$, donc à $\mathbf{v}$ et à $\mathbf{v}'$, vu que ces deux vecteurs appartiennent au plan vectoriel engendré par $\mathbf{u}$ et $\mathbf{u}'$.

2.8.11 Distance de deux droites gauches

Soit $\mathscr{S}$ et $\mathscr{S}'$ des sous-espaces affines de $\mathscr{E}$ de dimension finie. On appelle *distance de* $\mathscr{S}$ *et* $\mathscr{S}'$, et l'on note $\operatorname{dist}(\mathscr{S}, \mathscr{S}')$, le minimum de $\delta(P_0, P_0')$, P_0 et P_0' parcourant respectivement $\mathscr{S}$ et $\mathscr{S}'$. Nous nous limiterons à examiner le cas où $\mathscr{S}$ et $\mathscr{S}'$ sont deux droites gauches, que nous noterons $\mathscr{D}$ et $\mathscr{D}'$. Soit $\mathbf{n}$ un vecteur directeur de leur perpendiculaire commune, P et P' respectivement le point de $\mathscr{D}$

et le point de $\mathscr{D}'$ appartenant à cette perpendiculaire. Si P_0 est un point quelconque de $\mathscr{D}$ et P'_0 un point quelconque de $\mathscr{D}'$, alors

$$\operatorname{dist}(\mathscr{D}, \mathscr{D}') = \delta(P, P') = \| \operatorname{proj}_{\mathbf{n}} \overrightarrow{P_0P'_0} \| = \frac{|(\overrightarrow{P_0P'_0} \mid \mathbf{n})|}{\| \mathbf{n} \|}. \tag{2.62}$$

La vérification de cette formule est analogue à celle de la formule (2.56); c'est pourquoi nous la laissons comme exercice au lecteur.

Dans le cas particulier où la dimension de $\mathscr{E}$ est 3, on posera $\mathbf{n} = \mathbf{v} \times \mathbf{v}'$, où $\mathbf{v}$ est un vecteur directeur de $\mathscr{D}$ et $\mathbf{v}'$ un vecteur directeur de $\mathscr{D}'$.

2.8.12 Angle déterminé par trois points

Supposons que les points P et R soient distincts d'un point Q. On appelle *angle déterminé par P, Q, R*, et l'on note $P\hat{Q}R$, l'angle des vecteurs $\overrightarrow{QP}$ et $\overrightarrow{QR}$ (fig. 2.7).

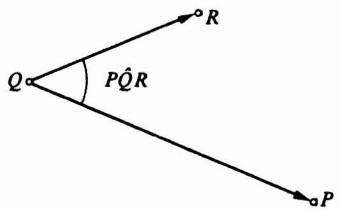

Fig. 2.7

2.8.13 Théorèmes de Pythagore et du cosinus

Soit P, Q, et R comme dans 2.8.12. D'après (2.24),

$$\begin{aligned}\| \overrightarrow{RP} \|^2 &= \| \overrightarrow{QP} - \overrightarrow{QR} \|^2 \\ &= \| \overrightarrow{QP} \|^2 + \| \overrightarrow{QR} \|^2 - 2 \| \overrightarrow{QP} \| \| \overrightarrow{QR} \| \cos(P\hat{Q}R),\end{aligned}$$

ce qui s'écrit aussi sous la forme

$$\delta^2(P, R) = \delta^2(P, Q) + \delta^2(Q, R) - 2\delta(P, Q)\delta(Q, R)\cos(P\hat{Q}R). \tag{2.63}$$

Dans le cas particulier où $P\hat{Q}R = \frac{\pi}{2}$, cela devient

$$\delta^2(P, R) = \delta^2(P, Q) + \delta^2(Q, R). \tag{2.64}$$

Les relations (2.63) et (2.64) sont appelées respectivement *théorèmes du cosinus* et *théorème de Pythagore*. On remarquera que (2.64) est encore vraie si $P = Q$ ou $R = Q$.

2.8.14 Angle d'une droite et un sous-espace affine

Soit $\mathscr{D}$ une droite de $\mathscr{E}$ de vecteur directeur $\mathbf{v}$ et $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et de dimension finie non nulle k. Si $\mathscr{D}$ n'est pas orthogonale à $\mathscr{S}$, on appelle *angle de $\mathscr{D}$ et $\mathscr{S}$* l'angle θ des vecteurs $\mathbf{v}$ et $\operatorname{proj}_S \mathbf{v}$ (fig. 2.8). En choisissant une base orthogonale quelconque $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ de S et en exprimant $\operatorname{proj}_S \mathbf{v}$ sous la forme (2.32), nous voyons aussitôt que $\cos\theta$ est positif, donc que θ appartient à l'intervalle $[0, \frac{\pi}{2}]$. Si $\mathscr{D}$ est orthogonale à $\mathscr{S}$, l'angle de $\mathscr{D}$ et $\mathscr{S}$ est, par convention, $\frac{\pi}{2}$ (c'est-à-dire l'angle de $\mathbf{v}$ et n'importe quel vecteur non nul de $\mathscr{S}$).

On remarquera que l'intersection de $\mathscr{D}$ et $\mathscr{S}$ peut être vide. En particulier, notre définition d'angle s'applique à deux droites gauches.

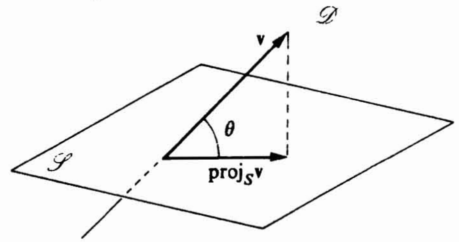

Fig. 2.8

2.8.15 Angle de deux hyperplans

Si $\mathscr{E}$ est de dimension finie supérieure à 1, on appelle *angle des hyperplans $\mathscr{S}$ et $\mathscr{S}'$* de $\mathscr{E}$ l'angle θ de deux droites $\mathscr{D}$ et $\mathscr{D}'$ orthogonales respectivement à $\mathscr{S}$ et à $\mathscr{S}'$ (fig. 2.9). Il est évident que $\cos\theta = \dfrac{|(\mathbf{n} \mid \mathbf{n}')|}{\| \mathbf{n} \| \| \mathbf{n}' \|}$, où $\mathbf{n}$ et $\mathbf{n}'$ sont des vecteurs normaux respectivement à $\mathscr{S}$ et à $\mathscr{S}'$. Il est également évident que lorsque $\mathscr{E}$ est de dimension 2, la présente définition de θ coïncide avec celle du paragraphe précédent.

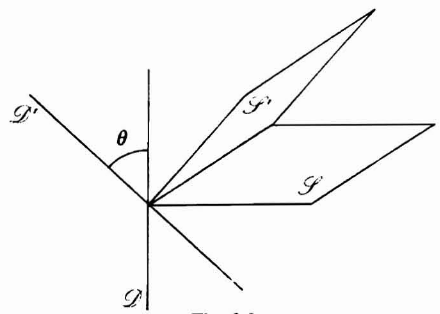

Fig. 2.9

2.8.16 Repères orthonormaux

Supposons que $\mathscr{E}$ soit de dimension finie non nulle n. On dit qu'un repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de $\mathscr{E}$ est *orthonormal* si la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormale.

2.8.17 Retour à l'équation cartésienne d'un hyperplan

Encore sous l'hypothèse que $\mathscr{E}$ est de dimension finie non nulle, considérons un hyperplan $\mathscr{S}$ de $\mathscr{E}$ de direction S, un vecteur $\mathbf{n}$ normal à $\mathscr{S}$ et un point quelconque P_0 de $\mathscr{S}$. D'après (1.16), P est un point de $\mathscr{S}$ si et seulement si le vecteur $\overrightarrow{P_0P}$ appartient à S. D'autre part, en vertu de (2.31), ce vecteur appartient à S si et seulement s'il est orthogonal à $\mathbf{n}$. L'hyperplan $\mathscr{S}$ est donc constitué des points P qui vérifient l'équation

$$(\mathbf{n} \mid \overrightarrow{P_0P}) = 0. \tag{2.65}$$

Munissons $\mathscr{E}$ d'un repère orthonormal $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et désignons par $x_1^0, x_2^0, ..., x_n^0$ les coordonnées de P_0, par $x_1, x_2, ..., x_n$ celles de P et par $v_1, v_2, ..., v_n$ les composantes de $\mathbf{n}$. L'équation (2.65) s'écrit, de manière équivalente, sous la forme

$$v_1(x_1 - x_1^0) + v_2(x_2 - x_2^0) + ... + v_n(x_n - x_n^0) = 0, \tag{2.66}$$

ou encore

$$v_1x_1 + v_2x_2 + ... + v_nx_n = \delta, \tag{2.67}$$

où

$$\delta = v_1x_1^0 + v_2x_2^0 + ... + v_nx_n^0 = (\mathbf{n} \mid \overrightarrow{OP_0}). \tag{2.68}$$

L'équation (2.67) n'est rien d'autre que l'équation cartésienne de l'hyperplan $\mathscr{S}$ (cf. 1.11.8). Elle est ici obtenue par un procédé propre aux espaces affines euclidiens. Ce procédé montre que les coefficients du premier membre de cette équation sont les composantes d'un vecteur normal à l'hyperplan. Grâce à (2.56), il fournit en outre une interprétation du second membre:

$$|\delta| = \|\mathbf{n}\| \operatorname{dist}(O, \mathscr{S}). \tag{2.69}$$

2.8.18 Problèmes de géométrie analytique euclidienne

La géométrie analytique euclidienne traite les problèmes de la géométrie euclidienne (distance, orthogonalité, angles, ...) par des calculs dans $\mathbb{R}^n$. Dans les exemples suivants, l'espace affine euclidien est supposé muni d'un repère orthonormal.

(1) Déterminer la distance du point $P(0, 1, -1, 2)$ à l'hyperplan $\mathscr{S}$ d'équation

$$x_1 - x_2 + 3x_3 + x_4 = 2.$$

Solution. En appliquant la formule (2.56) à $P(0, 1, -1, 2)$, $P_0(0, 0, 0, 2)$ et $\mathbf{n}(1, -1, 3, 1)$, nous obtenons

$$\operatorname{dist}(P, \mathscr{S}) = \frac{4}{\sqrt{12}} = \frac{2}{\sqrt{3}}.$$

(2) Calculer la distance des deux droites gauches $\mathscr{D}$ et $\mathscr{D}'$ d'équations paramétriques respectives

$$\begin{aligned} x_1 &= 1 & \qquad x_1 &= -1 - 2\alpha' \\ x_2 &= 1 + \alpha & x_2 &= 2 - \alpha' \\ x_3 &= 1 + 2\alpha, & x_3 &= -2\alpha'. \end{aligned}$$

Solution. $\mathbf{v}(0, 1, 2)$ est un vecteur directeur de $\mathscr{D}$ et $\mathbf{v}'(-2, -1, -2)$ un vecteur directeur de $\mathscr{D}'$. En appliquant la formule (2.62) à $P_0(1, 1, 1)$, $P_0'(-1, 2, 0)$ et $\mathbf{n} = \mathbf{v} \times \mathbf{v}'$, nous obtenons

$$\operatorname{dist}(\mathscr{D}, \mathscr{D}') = \frac{|-3|}{\sqrt{5}} = \frac{3}{\sqrt{5}}.$$

(3) Les données étant celles du problème précédent, déterminer les points d'intersection P, P' des droites $\mathscr{D}$, $\mathscr{D}'$ et la perpendiculaire commune à $\mathscr{D}$ et à $\mathscr{D}'$.

Solution. En appliquant les formules (2.61) à $P_0(1, 1, 1)$, $P_0'(-1, 2, 0)$, $\mathbf{v}(0, 1, 2)$ et $\mathbf{v}'(-2, -1, -2)$, nous obtenons

$$\alpha = \frac{4}{5}, \quad \alpha' = -1.$$

Les coordonnées des points cherchés P et P' se calculent alors par (2.57):

$$P(1, \frac{9}{5}, \frac{13}{5}), \quad P'(1, 3, 2).$$

(4) Trouver la projection orthogonale du point $P(-1, 1, 2, 3)$ sur le plan $\mathscr{S}$ passant par le point $P_0(0, -1, 0, -1)$ et de vecteurs directeurs $\mathbf{v}_1(1, 1, 0, 0)$ et $\mathbf{v}_2(1, 1, 1, 1)$. Déterminer, en outre, la distance de P à $\mathscr{S}$.

Solution. On voit aussitôt que $\mathbf{v}_1$ et $\mathbf{v}_2 - \mathbf{v}_1$ sont des vecteurs directeurs orthogonaux de $\mathscr{S}$. D'après 2.8.7 et (2.32),

$$\operatorname{proj}_{\mathscr{S}} P = P_0 + \operatorname{proj}_S \overrightarrow{P_0P} = P_0 + \frac{(\overrightarrow{P_0P} \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 + \frac{(\overrightarrow{P_0P} \mid \mathbf{v}_2 - \mathbf{v}_1)}{(\mathbf{v}_2 - \mathbf{v}_1 \mid \mathbf{v}_2 - \mathbf{v}_1)}(\mathbf{v}_2 - \mathbf{v}_1),$$

ce qui nous permet de calculer les coordonnées de $\operatorname{proj}_{\mathscr{S}} P$, soit $\frac{1}{2}$, $-\frac{1}{2}$, 3, 2.

La distance de P à $\mathscr{S}$ est égale à la norme du vecteur $\overrightarrow{P(\operatorname{proj}_{\mathscr{S}} P)}$, c'est-à-dire à $\sqrt{13/2}$.

(5) Trouver l'équation du lieu géométrique des points équidistants des points $P_1(1, 0, 2, -1)$ et $P_2(0, 1, -1, 1)$.

Solution. Un point $P(x_1, x_2, x_3, x_4)$ est situé à égale distance de P_1 et de P_2 si et seulement si

$$(x_1 - 1)^2 + x_2^2 + (x_3 - 2)^2 + (x_4 + 1)^2 = x_1^2 + (x_2 - 1)^2 + (x_3 + 1)^2 + (x_4 - 1)^2.$$

En développant les carrés, on voit aisément que cette équation se réduit à

$$2x_1 - 2x_2 + 6x_3 - 4x_4 = 3.$$

C'est l'équation d'un hyperplan, appelé *hyperplan médiateur* du segment de droite P_1P_2.

(6) Déterminer l'équation du cône de révolution de sommet $P(3, 0, 3)$ et dont l'intersection avec le plan $\mathscr{S}$ d'équation

$$2x_1 - x_2 + 2x_3 = 3$$

est un cercle de rayon 3.

Solution. En appliquant la formule (2.56) à $P(3, 0, 3)$, $P_0(0, -3, 0)$ et $\mathbf{n}(2, -1, 2)$, nous obtenons

$$\operatorname{dist}(P, \mathscr{S}) = \frac{|-9|}{3} = 3.$$

La tangente de l'angle θ des génératrices et l'axe du cône est donc

$$\operatorname{tg}\theta = \frac{3}{3} = 1,$$

ce qui entraîne $\theta = \frac{\pi}{4}$. Soit $Q(x_1, x_2, x_3)$ un point quelconque du cône distinct de P. L'angle du vecteur directeur $\mathbf{n}(2, -1, 2)$ de l'axe du cône et $\overrightarrow{PQ}(x_1 - 3, x_2, x_3 - 3)$ étant $\frac{\pi}{4}$ ou $\frac{3\pi}{4}$, il s'ensuit, grâce à (2.23), que

$$2(x_1 - 3) - x_2 + 2(x_3 - 3) = ((x_1 - 3)^2 + x_2^2 + (x_3 - 3)^2)^{1/2} \cdot 3 \cdot (\pm\frac{\sqrt{2}}{2}).$$

On notera que les coordonnées de P vérifient également cette équation. Il ne reste alors plus qu'à élever les deux membres au carré et à développer ces carrés pour obtenir l'équation demandée:

$$x_1^2 + 7x_2^2 + x_3^2 + 8x_1x_2 - 16x_1x_3 + 8x_2x_3 + 42x_1 - 48x_2 + 42x_3 = 126.$$

2.9 EXERCICES

2.9.1 Lesquelles des opérations définies ci-dessous sont des produits scalaires dans $\mathbb{R}^3$?

(a) $(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 + x_2y_3 + x_3y_2 + 2x_3y_3$.
(b) $(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 - 3x_2y_3 + x_3y_2 + x_3y_3$.
(c) $(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 + x_3^2y_2$.

Même question rapportée aux espaces vectoriels respectifs $C_{[-1,1]}$, $C_{[0,\infty)}$ et $C_{\mathbb{R}}$:

(d) $(\mathbf{f} \mid \mathbf{g}) = \int_{-1}^{1} t^2\mathbf{f}(t)\mathbf{g}(t)\mathrm{d}t$.

(e) $(\mathbf{f} \mid \mathbf{g}) = \int_{0}^{\infty} t\mathbf{f}^2(t)\mathbf{g}(t)\mathrm{d}t$.

(f) $(\mathbf{f} \mid \mathbf{g}) = \int_{-\infty}^{\infty} e^{-t}\mathbf{f}(t)\mathbf{g}(t)\mathrm{d}t$.

2.9.2 Pour tout couple de polynômes $\mathbf{p}(t) \equiv \sum_{i=0}^{m} \alpha_i t^i$ et $\mathbf{q}(t) \equiv \sum_{i=0}^{n} \beta_i t^i$, posons

$$(\mathbf{p} \mid \mathbf{q}) = \sum_{i=0}^{\min(m,n)} \alpha_i\beta_i.$$

(a) Montrer que l'opération ainsi définie est un produit scalaire dans l'espace vectoriel des polynômes.
(b) Exhiber une famille infinie de polynômes orthonormale pour ce produit scalaire.

2.9.3 Montrer que pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs d'un espace vectoriel euclidien,

$$\| \mathbf{x} + \mathbf{y} \|^2 + \| \mathbf{x} - \mathbf{y} \|^2 = 2\| \mathbf{x} \|^2 + 2\| \mathbf{y} \|^2$$ (*identité du parallélogramme*).

2.9.4 Montrer que dans un espace vectoriel euclidien la somme et la différence de deux vecteurs de même norme sont orthogonales.

2.9.5 Par le procédé de Gram-Schmidt, orthogonaliser le triplet de vecteurs de $\mathbb{R}^4$

$$\left(\begin{pmatrix}1\\0\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\-1\end{pmatrix}, \begin{pmatrix}0\\3\\-1\\0\end{pmatrix}\right).$$

2.9.6 Soit $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ une base orthonormale d'un espace vectoriel euclidien de dimension 4. Par le procédé de Gram-Schmidt, orthogonaliser le triplet de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, où $\mathbf{x}_1 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4$, $\mathbf{x}_2 = \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4$ et $\mathbf{x}_3 = \mathbf{e}_3 + \mathbf{e}_4$.

2.9.7 A l'aide de l'inégalité de Cauchy-Schwarz, montrer que

$$\int_0^1 \sqrt{1 + t^3}\,\mathrm{d}t < \frac{\sqrt{5}}{2}.$$

2.9.8 Exprimer le vecteur-colonne $\begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$ sous la forme d'une combinaison linéaire des vecteurs-colonnes

$$\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix},\ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},\ \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}.$$

2.9.9 Soit $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k)$ une famille orthonormale de vecteurs d'un espace vectoriel euclidien E. Montrer que pour tout vecteur $\mathbf{x}$ de E,

$$\sum_{i=1}^{k} (\mathbf{x} \mid \mathbf{e}_i)^2 \leqslant \|\mathbf{x}\|^2.$$

Dans quel cas les deux membres sont-ils égaux?

2.9.10 Soit E un espace vectoriel euclidien de dimension 4, muni d'une base orthonormale. Déterminer la projection orthogonale du vecteur $\mathbf{x}(3, 2, -2, -1)$ sur le plan vectoriel de E engendré par les vecteurs $\mathbf{v}_1(1, 0, -1, 1)$ et $\mathbf{v}_2(1, 0, 0, 1)$.

2.9.11 Soit E un espace vectoriel euclidien de dimension 4, muni d'une base orthonormale. Soit S le sous-espace vectoriel de E engendré par les vecteurs $\mathbf{v}_1(1, 1, 1, 1)$ $\mathbf{v}_2(-3, 1, 1, 1)$ et $\mathbf{v}_3(0, -2, 1, 1)$. Déterminer le complémentaire orthogonal de S dans E.

2.9.12 Calculer la meilleure approximation de la fonction $\mathbf{f}(t) \equiv t^2$ de $C_{[-\pi, \pi]}$ par des polynômes trigonométriques d'ordre inférieur ou égal à k.

2.9.13 Soit S le sous-espace vectoriel de $C_{[-\pi, \pi]}$ formé de tous les polynômes trigonométriques.

(a) A l'aide de (2.34), montrer que $S^\perp = \{\mathbf{0}\}$. (Prendre un $\mathbf{f}$ de $S^\perp$ et s'assurer que $\mathbf{f}_k = \mathbf{0}$.)

(b) En déduire que S n'admet aucun complémentaire orthogonal dans $C_{[-\pi, \pi]}$ (autrement dit, que $S \oplus S^\perp \neq C_{[-\pi, \pi]}$).

2.9.14 Soit E un espace vectoriel euclidien, orienté et de dimension 3. A quelle condition doivent satisfaire deux vecteurs donnés $\mathbf{y}$ et $\mathbf{z}$ de E, pour que l'équation vectorielle $\mathbf{x} \times \mathbf{y} = \mathbf{z}$ ait au moins une solution $\mathbf{x}_0$ dans E? Quel est alors l'ensemble des solutions?

2.9.15 Soit E comme dans l'exercice précédent. Démontrer les identités vectorielles suivantes:

(a) $(\mathbf{x} \times (\mathbf{y} \times \mathbf{z})) + (\mathbf{y} \times (\mathbf{z} \times \mathbf{x})) + (\mathbf{z} \times (\mathbf{x} \times \mathbf{y})) = \mathbf{0}$. (Utiliser (2.50).)

(b) $(\mathbf{x} \times \mathbf{y} \mid (\mathbf{y} \times \mathbf{z}) \times (\mathbf{z} \times \mathbf{x})) = [\mathbf{x}, \mathbf{y}, \mathbf{z}]^2$. (Utiliser (2.51) ou (2.52).)

(c) $\mathbf{v} = \dfrac{[\mathbf{v}, \mathbf{y}, \mathbf{z}]}{[\mathbf{x}, \mathbf{y}, \mathbf{z}]}\mathbf{x} + \dfrac{[\mathbf{x}, \mathbf{v}, \mathbf{z}]}{[\mathbf{x}, \mathbf{y}, \mathbf{z}]}\mathbf{y} + \dfrac{[\mathbf{x}, \mathbf{y}, \mathbf{v}]}{[\mathbf{x}, \mathbf{y}, \mathbf{z}]}\mathbf{z}$ $([\mathbf{x}, \mathbf{y}, \mathbf{z}] \neq 0)$. (Utiliser (2.52).)

Exercices sur les espaces affines euclidiens

Dans les exercices suivants, lorsqu'il est sous-entendu que l'espace affine euclidien est muni d'un repère, celui-ci est supposé orthonormal.

2.9.16 Calculer la distance du point $P(0, 1, -1, 2)$ à l'hyperplan d'équation $x_1 - x_2 + 3x_3 + x_4 = 2$.

2.9.17 Trouver l'équation de l'hyperplan médiateur du segment d'extrémités $P_1(3, -2, 5, -1)$ et $P_2(-3, 1, 2, 4)$.

2.9.18 Trouver les équations paramétriques des bissectrices des deux droites passant par le point $P_0(-1, 2, 1, 3)$, de vecteurs directeurs respectifs $\mathbf{v}(1, -3, 2, 2)$ et $\mathbf{v}'(4, 0, 1, 1)$.

2.9.19 Déterminer la projection orthogonale du point $P(-1, 1, 1, 1)$ sur le plan passant par le point $P_0(0, 1, 2, -1)$, de vecteurs directeurs $\mathbf{v}_1(-1, 0, 2, 1)$ et $\mathbf{v}_2(0, 1, 1, -1)$. Calculer en outre la distance de P à ce plan.

2.9.20 Dans un espace affine euclidien, orienté et de dimension 3, on considère un point P et la droite $\mathscr{D}$ passant par le point P_0 et de vecteur directeur $\mathbf{v}$. Montrer que

$$\operatorname{dist}(P, \mathscr{D}) = \frac{\| \overrightarrow{P_0P} \times \mathbf{v} \|}{\| \mathbf{v} \|}.$$

2.9.21 Déterminer la perpendiculaire commune et la distance des deux droites d'équations paramétriques respectives

$$\begin{aligned} x_1 &= \alpha \\ x_2 &= 1 \\ x_3 &= -1 + \alpha \\ x_4 &= 2 - \alpha, \end{aligned} \qquad \begin{aligned} x_1 &= \alpha' \\ x_2 &= 1 - \alpha' \\ x_3 &= 0 \\ x_4 &= -1 + 2\alpha'. \end{aligned}$$

2.9.22 Calculer l'angle de la droite passant par le point $P_0(1, 1, 1, 1)$, de vecteur directeur $\mathbf{v}(0, 1, -1, 0)$, et le plan passant par le point $Q_0(1, 0, -1, 1)$, de vecteurs directeurs $\mathbf{v}_1(1, 1, 0, 0)$ et $\mathbf{v}_2(0, 1, 1, -1)$.

2.9.23 Déterminer l'angle que forment la droite et l'hyperplan de l'exercice 1.12.18.

2.9.24 Déterminer l'angle des deux hyperplans d'équations respectives $x_1 - 3x_2 + x_3 - 2x_4 - x_5 = 2$ et $2x_1 - 2x_2 - 3x_3 + 2x_4 + 2x_5 = -1$.

2.9.25 Soit $\mathscr{E}$ un espace affine euclidien de dimension finie supérieure à 1.

(a) Montrer que les équations des hyperplans bissecteurs des deux hyperplans de $\mathscr{E}$ d'équations respectives $(\mathbf{n}_1 \mid \mathbf{x}) = \delta_1$ et $(\mathbf{n}_2 \mid \mathbf{x}) = \delta_2$ sont données par

$$\frac{(\mathbf{n}_1 \mid \mathbf{x}) - \delta_1}{\| \mathbf{n}_1 \|} = \pm \frac{(\mathbf{n}_2 \mid \mathbf{x}) - \delta_2}{\| \mathbf{n}_2 \|}.$$

(b) Appliquer ce résultat à $\mathbf{n}_1(1, -1, 3, 1, 2)$, $\mathbf{n}_2(-1, 2, 4, 0, -2)$, $\delta_1 = -1$ et $\delta_2 = 2$.

2.9.26 Dans un espace affine euclidien, orienté et de dimension 3, on considère trois points P_1, P_2, P_3 non alignés et on pose $\theta_1 = P_3\hat{P}_1P_2$, $\theta_2 = P_1\hat{P}_2P_3$ et $\theta_3 = P_2\hat{P}_3P_1$. A l'aide du produit vectoriel, démontrer que

$$\frac{\sin\theta_1}{\delta(P_2, P_3)} = \frac{\sin\theta_2}{\delta(P_1, P_3)} = \frac{\sin\theta_3}{\delta(P_1, P_2)} \quad (\textit{théorème du sinus}).$$

2.9.27 Les quatre sommets d'un tétraèdre sont $P_1(1, 1, 1)$, $P_2(-1, 2, 1)$, $P_3(5, 2, 3)$ et P_4. Trouver les coordonnées du point P_4, sachant qu'il appartient à la droite passant par $P_0(1, 0, 1)$, de vecteur directeur $\mathbf{v}(-2, -1, 3)$, et que le volume du tétraèdre est 5.

CHAPITRE 3

Systèmes linéaires

3.1 DÉFINITIONS ET EXEMPLES

3.1.1 Introduction

Nous avons déjà eu l'occasion de constater que la résolution de certains problèmes liés aux notions de sous-espace vectoriel et de sous-espace affine se réduit à la recherche des solutions d'un système d'équations linéaires. D'un point de vue plus général, on voit souvent que des problèmes d'origines les plus diverses se traduisent en termes d'équations linéaires. Le présent chapitre exposera les principaux résultats de la théorie des systèmes d'équations linéaires. L'idée à la base de notre étude est simple: par une succession d'opérations élémentaires sur les équations, le système est transformé en un système échelonné équivalent que l'on peut résoudre facilement. Le procédé de transformation est connu sous le nom de *méthode de Gauss* ou *méthode d'élimination des inconnues.*

3.1.2 Systèmes linéaires

On appelle *système linéaire*, ou simplement *système*, toute famille d'équations de la forme

$$\begin{aligned} a_{11}x_1 &+ a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 &+ a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ &\dots\dots \\ a_{m1}x_1 &+ a_{m2}x_2 + \dots + a_{mn}x_n = b_m, \end{aligned} \tag{3.1}$$

où $a_{11}, a_{12}, \dots, a_{mn}$ sont des nombres appelés *coefficients du système*, $b_1, b_2, \dots, b_m$ des nombres appelés *coefficients du second membre* et $x_1, x_2, \dots, x_n$ des nombres inconnus appelés *inconnues* du système. Si les coefficients du second membre sont tous nuls, on dit que le système est *homogène*. On appelle *système homogène associé* au système (3.1) le système que l'on obtient de (3.1) en substituant des zéros aux coefficients du second membre.

3.1.3 Ecriture abrégée

On écrit souvent le système (3.1) sous la forme abrégée suivante:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \dots, m. \tag{3.2}$$

Une écriture encore plus condensée sera introduite dans l'exemple (5) de 4.1.7.

3.1.4 Solutions d'un système

On appelle *solution* du système (3.1) tout n-uplet de nombres $(x_1^0, x_2^0, ..., x_n^0)$ tel que

$$\sum_{j=1}^{n} a_{ij}x_j^0 = b_i, \quad i = 1, 2, ..., m.$$

Résoudre un système signifie trouver l'ensemble des solutions de ce système. Deux systèmes à n inconnues sont dits *équivalents* si toute solution de l'un est solution de l'autre, autrement dit, s'ils admettent le même ensemble de solutions. On dit parfois que les équations d'un système sont *compatibles* ou *incompatibles*, suivant que ce système admet au moins une solution ou n'en admet aucune.

3.1.5 Interprétation géométrique

Supposons que les premiers membres des équations du système (3.1) soient non nuls. D'après 1.11.8, chacune de ces équations représente alors un hyperplan d'un espace affine de dimension n. Par conséquent, l'ensemble des solutions du système, regardé comme ensemble de n-uplets de coordonnées, représente une intersection finie d'hyperplans. Selon 1.10.7, une telle intersection est un sous-espace affine ou l'ensemble vide. Du point de vue géométrique, résoudre un système admettant au moins une solution revient à chercher les équations paramétriques (1.19) de ce sous-espace affine.

3.1.6 Exemples

(1) Le système

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 0 \end{aligned}$$

n'admet aucune solution, car les deux équations se contredisent l'une l'autre. Ces équations représentent deux droites parallèles.

(2) Nous nous proposons de résoudre le système

$$\begin{aligned} 3x_1 - 3x_2 - 2x_3 &= -1 \\ 4x_1 - x_2 + 3x_3 &= -1 \\ -2x_1 + 4x_2 + 3x_3 &= 3 \end{aligned} \tag{3.3}$$

par la méthode d'élimination des inconnues.

A l'aide de la première équation, exprimons x_1 en fonction de x_2 et x_3:

$$x_1 = \tfrac{1}{3}(-1 + 3x_2 + 2x_3). \tag{3.4}$$

Substituons ensuite l'expression ainsi obtenue à x_1 dans les deux autres équations de (3.3):

$$\begin{aligned} \tfrac{4}{3}(-1 + 3x_2 + 2x_3) - x_2 + 3x_3 &= -1 \\ -\tfrac{2}{3}(-1 + 3x_2 + 2x_3) + 4x_2 + 3x_3 &= 3. \end{aligned}$$

En multipliant chaque membre par 3 et en groupant les termes, nous obtenons

$$\begin{aligned} 9x_2 + 17x_3 &= 1 \\ 6x_2 + 5x_3 &= 7. \end{aligned} \tag{3.5}$$

A l'aide de la première de ces deux équations, exprimons à présent x_2 en fonction de x_3 et substituons l'expression obtenue à x_2 dans la deuxième équation. Après simplification, il reste

$$x_3 = -1 \tag{3.6}$$

Par substitution de cette valeur à x_3 dans l'une des deux équations (3.5), nous déterminons la valeur de x_2, soit $x_2 = 2$. Nous calculons finalement la valeur de x_1 par (3.4), ce qui nous donne $x_1 = 1$. Le système (3.3) admet donc l'unique solution $x_1 = 1$, $x_2 = 2$, $x_3 = -1$.

Les deux équations (3.5) s'obtiennent plus rapidement en soustrayant la première équation de (3.3) multipliée par 4 de la deuxième multipliée par 3 et en additionnant la même équation multipliée par 2 à la troisième multipliée par 3. De même, nous obtenons l'équation (3.6) en soustrayant la première équation de (3.5) multipliée par 2 de la deuxième multipliée par 3, puis en divisant les deux membres par -19. Ces opérations réduisent le système (3.3) au système équivalent

$$\begin{aligned} 3x_1 - 3x_2 - 2x_3 &= -1 \\ 9x_2 + 17x_3 &= 1 \\ x_3 &= -1 \end{aligned} \tag{3.7}$$

que l'on résout par substitution, comme il a été indiqué ci-dessus. Par la suite, c'est cette variante de la méthode d'élimination des inconnues qui retiendra notre attention.

(3) Nous allons maintenant résoudre le système

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ 3x_1 + 2x_2 - 3x_3 &= 2. \end{aligned} \tag{3.8}$$

Eliminons l'inconnue x_1 de la deuxième équation en soustrayant la première équation multipliée par 3 de la deuxième. Cette opération réduit le système (3.8) au système équivalent

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ 5x_2 - 9x_3 &= 2. \end{aligned} \tag{3.9}$$

Pour résoudre ce système, nous assignons à x_3 une valeur arbitraire α, exprimons x_2 en fonction de α à l'aide de la deuxième équation, puis x_1 à l'aide de la première équation. Il en résulte que les solutions du système (3.8) sont les triplets de nombres de la forme

$$\begin{aligned} x_1 &= \tfrac{2}{5} - \tfrac{1}{5}\alpha \\ x_2 &= \tfrac{2}{5} + \tfrac{9}{5}\alpha \\ x_3 &= \alpha, \end{aligned} \tag{3.10}$$

où α parcourt $\mathbb{R}$. L'ensemble des solutions représente donc une droite.

3.2 EXISTENCE ET UNICITÉ DES SOLUTIONS

3.2.1 Matrices

On appelle *matrice* à m lignes et n colonnes, ou matrice de type $m \times n$, tout tableau rectangulaire de nombres

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ & & \cdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ & & \cdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}. \tag{3.11}$$

On désigne souvent une matrice de type $m \times n$ plus brièvement par $(a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ ou simplement par (a_{ij}).

Le nombre a_{ij} est appelé *terme d'indices i, j*. L'indice i est appelé *indice de ligne* et l'indice j *indice de colonne*.

Lorsque $m = n$, on dit que (a_{ij}) est une *matrice carrée* d'ordre n. Dans ce cas, les termes $a_{11}, a_{22}, \ldots, a_{nn}$ sont appelés *termes diagonaux*.

On appelle une matrice à une seule ligne *matrice-ligne* et une matrice à une seule colonne *matrice-colonne*. Il est clair qu'une matrice-colonne n'est rien d'autre qu'un vecteur-colonne. Par la suite, les lignes d'une matrice seront assimilées à des matrices-lignes et les colonnes à des matrices-colonnes.

L'intérêt de la notion de matrice apparaîtra tout au long de ce livre, à partir de cette section, mais la raison d'être immédiate de cette notion est simplement de permettre à certaines familles finies de nombres d'être conçues sous la forme d'un tableau rectangulaire.

3.2.2 Notation

Nous assignerons aux matrices des symboles propres, à savoir les lettres latines majuscules imprimées en caractère gras: **A**, **B**, ... Toutefois, les matrices-colonnes seront également désignées par les symboles réservés aux vecteurs: **a**, **b**, ... ; nous les appellerons d'ailleurs indifféremment matrices-colonnes ou vecteurs-colonnes.

3.2.3 Matrices nulles et matrices-unités

On appelle *matrice nulle*, et on note **O**, toute matrice dont chaque terme est nul. Les matrices-colonnes nulles sont également désignées par le symbole vectoriel **0**.

On appelle *matrice-unité* d'ordre n, et on note $\mathbf{I}_n$ ou simplement **I**, la matrice carrée d'ordre n

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & \cdots & & \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{3.12}$$

Nous verrons au chapitre suivant que **O** joue le rôle d'élément neutre de l'addition matricielle et **I** d'élément neutre de la multiplication matricielle.

3.2.4 Rang d'une matrice

Soit $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ les colonnes d'une matrice **A**. On appelle *rang* de **A**, et on note rg**A**, le rang de la famille $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ (cf. 1.8.2).

Par exemple, le rang de **O** est 0 et le rang de $\mathbf{I}_n$ est n; les rangs des matrices

$$\begin{pmatrix} 2 & 2 & 5 & 2 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 4 & -1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 2 & 5 \\ -1 & 3 & 1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 3 \\ -2 & 0 & -6 \\ 1 & 0 & 3 \end{pmatrix}$$

sont respectivement 3, 2 et 1.

3.2.5 Matrices associées à un système linéaire

On appelle *matrice associée au système* (3.1) la matrice (3.11), c'est-à-dire la matrice **A** dont les termes sont les coefficients du système. On appelle *matrice du second membre du système* (3.1), ou simplement *second membre du système* (3.1), la matrice-colonne $\mathbf{b} = (b_i)$ dont les termes sont les coefficients du second membre de ce système. On appelle *matrice augmentée associée au système* (3.1) la matrice obtenue de **A** en y ajoutant **b** comme $(n + 1)$-ième colonne.

Nous venons de définir les différentes matrices associées à un système, mais il est également utile de pouvoir parler du système homogène associé à une matrice donnée ou du système associé à une matrice augmentée donnée. Cette terminologie s'explique d'elle-même et nous ne nous attarderons donc pas à la préciser davantage.

3.2.6 Ecriture vectorielle d'un système linéaire

Considérons un système de matrice associée $\mathbf{A}$ et de second membre $\mathbf{b}$. Désignons par $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ les colonnes de $\mathbf{A}$. Le système s'écrit alors de manière équivalente sous la forme d'une équation vectorielle linéaire:

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n = \mathbf{b}. \tag{3.13}$$

En effet, les deux membres de cette équation sont les vecteurs-colonnes dont les i-ièmes termes sont les deux membres de la i-ième équation du système.

D'après la proposition 1.8.4, l'équation (3.13) admet au moins une solution $(x_1^0, x_2^0, \ldots, x_n^0)$ si et seulement si le rang de la famille $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ est égal au rang de la famille augmentée $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n, \mathbf{b})$. Selon la proposition 1.5.6, cette solution est unique si et seulement si le rang de la famille $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ est n.

A l'aide de la définition 3.2.4, nos conclusions se résument de la manière suivante:

3.2.7 Proposition. Existence et unicité des solutions d'un système linéaire

Pour qu'un système linéaire de matrice associée $\mathbf{A}$ *et de second membre* $\mathbf{b}$ *admette au moins une solution, il faut et il suffit que le rang de* $\mathbf{A}$ *soit égal au rang de la matrice augmentée* $(\mathbf{A}\,|\,\mathbf{b})$. *Si cette condition est remplie, le système admet une seule solution si et seulement si le rang de* $\mathbf{A}$ *est égal au nombre d'inconnues, autrement dit, les colonnes de* $\mathbf{A}$ *sont linéairement indépendantes.*

3.3 MATRICES ÉCHELONNÉES

3.3.1 Exemple

La méthode employée pour réduire les systèmes (3.3) et (3.8) aux systèmes équivalents (3.7) et (3.9) est, en fait, un procédé d'annulation de certains termes de la matrice associée au système par des opérations sur les lignes de la matrice augmentée. Voici, indiquée à côté de cette matrice, la suite des opérations que nous avons effectuées pour aboutir au système (3.7):

$$\begin{array}{l} \begin{pmatrix} 3 & -3 & -2 & -1 \\ 4 & -1 & 3 & -1 \\ -2 & 4 & 3 & 3 \end{pmatrix} \begin{array}{l} \\ 3\mathbf{L}_2 - 4\mathbf{L}_1 \\ 3\mathbf{L}_3 + 2\mathbf{L}_1 \end{array} \end{array} \tag{3.14}$$

$$\begin{pmatrix} 3 & -3 & -2 & -1 \\ 0 & 9 & 17 & 1 \\ 0 & 6 & 5 & 7 \end{pmatrix} \begin{array}{l} \\ \\ -\tfrac{1}{19}(3\mathbf{L}_3 - 2\mathbf{L}_2) \end{array}$$

$$\begin{pmatrix} 3 & -3 & -2 & -1 \\ 0 & 9 & 17 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}. \tag{3.15}$$

Le système associé à la matrice augmentée (3.15) est le système (3.7). Suivant l'expression que nous allons introduire dans 3.3.5, la matrice (3.14) a été réduite à la forme échelonnée (3.15) par des opérations élémentaires sur les lignes. Il est peut-être superflu de préciser que le symbole $\mathbf{L}_i$ utilisé ci-dessus désigne la i-ième ligne de la matrice contiguë et que la formule suivant la i-ième ligne indique les opérations à effectuer pour obtenir la i-ième ligne de la matrice successive.

3.3.2 Matrices échelonnées

On dit qu'une matrice est *échelonnée* si ses lignes satisfont aux deux conditions suivantes:

(a) Toute ligne nulle n'est suivie que de lignes nulles.

(b) L'indice de colonne du premier terme non nul de toute ligne non nulle est supérieur à l'indice de colonne du premier terme non nul de la ligne qui la précède.

Une matrice échelonnée non nulle est donc de la forme

$$\begin{pmatrix} a_{1j_1} & \cdots & & & & \cdots & a_{1n} \\ & & a_{2j_2} & \cdots & & \cdots & a_{2n} \\ & & & \cdots\cdots\cdots & & & \\ & 0 & & & a_{rj_r} & \cdots & a_{rn} \\ & & & & & & \end{pmatrix}, \tag{3.16}$$

où $j_1 < j_2 < \ldots < j_r$ et $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ sont des termes non nuls. Bien entendu, les lignes nulles terminales peuvent manquer.

Les matrices nulles et les matrices-unités $\mathbf{I}_n$ sont des matrices échelonnées. Deux autres exemples de matrices échelonnées sont les suivants:

$$\begin{pmatrix} 2 & 0 & -2 & 1 & 2 & -1 \\ 0 & 0 & 0 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 2 & -3 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

3.3.3 Rang d'une matrice échelonnée

Les colonnes d'indice $j_1\, j_2, \ldots, j_r$ de la matrice (3.16) sont clairement linéairement indépendantes. Envisagées comme des vecteurs-colonnes de $\mathbb{R}^r$, elles forment donc une base de cet espace vectoriel. En considérant les autres colonnes également comme des vecteurs-colonnes de $\mathbb{R}^r$, nous en déduisons qu'elles sont combinaison linéaire de celles d'indice $j_1, j_2, \ldots, j_r$ et donc que le rang de la matrice (3.16) est r.

On notera que r est le nombre de lignes non nulles de la matrice (3.16) et également le rang de la famille des lignes de cette matrice, puisque les lignes non nulles sont manifestement linéairement indépendantes.

3.3.4 Opérations élémentaires sur les lignes d'une matrice

On appelle *opération élémentaire sur les lignes d'une matrice* toute opération de l'un des trois types suivants:

Type 1: échanger deux lignes.
Type 2: additionner à une ligne un multiple d'une autre ligne.
Type 3: multiplier une ligne par un nombre non nul.

3.3.5 Réduction à la forme échelonnée

Toute matrice peut être *transformée en une matrice échelonnée* par une suite finie d'opérations de type 1 et 2. On dit également que toute matrice peut être *réduite à la forme échelonnée* par une telle suite d'opérations. Soit en effet $\mathbf{A} = (a_{ij})$ une matrice de type $m \times n$ non nulle. Si $m = 1$, $\mathbf{A}$ est échelonnée. Si $m > 1$, désignons par j_1 le plus petit indice de colonne non nulle. En échangeant deux lignes, si nécessaire, nous transformons $\mathbf{A}$ en une matrice $\mathbf{A}^{(1)} = (a^{(1)}_{ij})$ dont le terme $a^{(1)}_{1j_1}$ est non nul. En additionnant ensuite à chaque ligne d'indice $i \geqslant 2$ la première ligne multipliée par $-a^{(1)}_{ij_1}/a^{(1)}_{1j_1}$, nous voyons que $\mathbf{A}^{(1)}$ se réduit soit à une matrice de la forme

$$\begin{pmatrix} a^{(1)}_{1j_1} & \cdots & & & \cdots & a^{(1)}_{1n} \\ & & a'_{2j_2} & \cdots & \cdots & a'_{2n} \\ & 0 & & \cdots & & \\ & & a'_{mj_2} & \cdots & \cdots & a'_{mn} \end{pmatrix},$$

où les termes $a'_{2j_2}, a'_{3j_2}, \ldots, a'_{mj_2}$ ne sont pas tous nuls, soit à une matrice dont les lignes d'indice $i \geqslant 2$ sont nulles. Dans le deuxième cas, la matrice est échelonnée. Dans le premier cas, le même procédé de réduction appliqué à la matrice

$$\begin{pmatrix} & a'_{2j_2} & \cdots & \cdots & a'_{2n} \\ 0 & & \cdots & & \\ & a'_{mj_2} & \cdots & \cdots & a'_{mn} \end{pmatrix}$$

fournit une matrice $\mathbf{A}^{(2)}$, puis une matrice $\mathbf{A}^{(3)}$ et, ainsi de suite, jusqu'à $\mathbf{A}^{(m)}$ ou jusqu'à ce que la seule ligne non nulle d'une matrice $\mathbf{A}^{(r)}$ soit la première. A ce point, la matrice transformée est de la forme

$$\begin{pmatrix} a^{(1)}_{1j_1} & \cdots & & & & \cdots & a^{(1)}_{1n} \\ & & a^{(2)}_{2j_2} & \cdots & & \cdots & a^{(2)}_{2n} \\ & \cdot & & \cdots & & & \\ & 0 & & & a^{(r)}_{rj_r} & \cdots & a^{(r)}_{rn} \\ & & & & & & \end{pmatrix}. \qquad (3.17)$$

3.3.6 Pivots

Les termes $a^{(1)}_{1j_1}, a^{(2)}_{2j_2}, \ldots, a^{(r)}_{rj_r}$ de la matrice (3.17) sont appelés *pivots*. Dans les opérations de réduction, les pivots servent à annuler certains termes des colonnes dont ils font partie. Par extension, on appelle *pivot d'une matrice échelonnée* non nulle tout premier terme non nul d'une ligne non nulle. Les pivots de la matrice (3.16) sont donc $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$. D'après 3.3.3, le rang d'une matrice échelonnée est égal au nombre de pivots qu'elle comprend.

3.3.7 Réduction à la forme échelonnée simplifiée

Le procédé employé pour parvenir à la matrice échelonnée (3.17) peut être poursuivi. Plus précisément, chaque pivot peut être utilisé pour annuler tous les autres termes de la colonne dont il fait partie. Par exemple, nous annulons le terme $a^{(1)}_{1j_2}$ en additionnant à la première ligne la deuxième multipliée par $-a^{(1)}_{1j_2}/a^{(2)}_{2j_2}$. Après cette suite d'opérations, nous divisons encore chaque ligne non nulle par son pivot. Nous obtenons ainsi une matrice échelonnée dont les pivots sont des 1 et les colonnes dont ils font partie des vecteurs-colonnes de la base canonique de $\mathbb{R}^m$. Une telle matrice est appelée *matrice échelonnée simplifiée*. Ainsi, toute matrice non nulle peut être *transformée en une matrice échelonnée simplifiée* par une suite finie d'opérations élémentaires sur les lignes. On dit également que toute matrice non nulle peut être *réduite à la forme échelonnée simplifiée* par une telle suite d'opérations.

Voici deux exemples de matrices échelonnées simplifiées:

$$\begin{bmatrix} 0 & 1 & 2 & 0 & 3 & 0 & -2 \\ 0 & 0 & 0 & 1 & -2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

3.3.8 Exemple de réduction

(1) Nous allons réduire la matrice suivante à la forme échelonnée:

$$\begin{bmatrix} 5 & 10 & 11 & -1 & 1 & -41 & 0 \\ -3 & -6 & -7 & 1 & -1 & 27 & 0 \\ 6 & 12 & 14 & -2 & 2 & -52 & 6 \\ 2 & 4 & 5 & -1 & 1 & -19 & 3 \end{bmatrix} \begin{array}{l} L_1 - 2L_4 \\ L_2 + L_4 \\ L_3 - 3L_4 \\ \end{array} \tag{3.18}$$

$$\begin{bmatrix} 1 & 2 & 1 & 1 & -1 & -3 & -6 \\ -1 & -2 & -2 & 0 & 0 & 8 & 3 \\ 0 & 0 & -1 & 1 & -1 & 5 & -3 \\ 2 & 4 & 5 & -1 & 1 & -19 & 3 \end{bmatrix} \begin{array}{l} \\ L_2 + L_1 \\ \\ L_4 - 2L_1 \end{array}$$

$$\begin{bmatrix} 1 & 2 & 1 & 1 & -1 & -3 & -6 \\ 0 & 0 & -1 & 1 & -1 & 5 & -3 \\ 0 & 0 & -1 & 1 & -1 & 5 & -3 \\ 0 & 0 & 3 & -3 & 3 & -13 & 15 \end{bmatrix} \begin{array}{ll} & \\ & \\ L_3 - L_2 & \text{puis échanger} \\ L_4 + 3L_2 & L_3 \text{ et } L_4 \end{array}$$

$$\begin{bmatrix} 1 & 2 & 1 & 1 & -1 & -3 & -6 \\ 0 & 0 & -1 & 1 & -1 & 5 & -3 \\ 0 & 0 & 0 & 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.19}$$

(2) Pour réduire la matrice (3.18) à la forme échelonnée simplifiée, nous multiplions d'abord la deuxième ligne de (3.19) par -1 et divisons la troisième par 2; ensuite nous procédons ainsi:

$$\begin{bmatrix} 1 & 2 & 1 & 1 & -1 & -3 & -6 \\ 0 & 0 & 1 & -1 & 1 & -5 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} L_1 - L_2 - 2L_3 \\ L_2 + 5L_3 \\ \\ \end{array}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 & -2 & 0 & -15 \\ 0 & 0 & 1 & -1 & 1 & 0 & 18 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.20}$$

3.3.9 Remarques

(1) D'après 3.3.5, la réduction d'une matrice à la forme échelonnée peut se faire sans l'emploi d'opérations de type 3. En pratique, toutefois, il est souvent plus commode d'utiliser également ce type d'opération, notamment pour éviter l'introduction de fractions, lorsque les termes de la matrice sont des entiers (cf. exemple 3.3.1).

(2) Le lecteur aura noté que la suite d'opérations réduisant la matrice (3.18) à la forme échelonnée (3.19) n'est pas celle qui est exposée dans 3.3.5. Des modifications du mécanisme de réduction analogues à celles que nous avons opérées sont souvent avantageuses et varient de cas en cas. Notre but évident était d'éviter l'apparition de fractions.

(3) Lorsqu'une ligne a été modifiée, il faut se garder d'utiliser encore sa forme non modifiée. Cela peut entraîner de graves erreurs, comme le montre l'exemple simple suivant:

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{array}{l} L_1 + L_2 \\ L_2 + L_1 \end{array} \qquad \begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix} \begin{array}{l} L_1 - L_2 \\ L_2 - L_1 \end{array} \qquad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

3.3.10 Conservation du rang

Les opérations élémentaires sur les lignes d'une matrice ne modifient pas son rang. En effet, si $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ sont les colonnes d'une matrice $\mathbf{A}$ et $\mathbf{a}'_1, \mathbf{a}'_2, \dots, \mathbf{a}'_n$ celles de la matrice $\mathbf{A}'$ déduite de $\mathbf{A}$ par une opération de type 1, 2 ou 3, alors toute relation de la forme

$$\alpha_1\mathbf{a}_{j_1} + \alpha_2\mathbf{a}_{j_2} + \dots + \alpha_k\mathbf{a}_{j_k} = \mathbf{0}$$

implique la relation

$$\alpha_1\mathbf{a}'_{j_1} + \alpha_2\mathbf{a}'_{j_2} + \dots + \alpha_k\mathbf{a}'_{j_k} = \mathbf{0}$$

et inversement, ce qui prouve qu'à toute famille libre de colonnes de $\mathbf{A}$ correspond une famille libre de colonnes de $\mathbf{A}'$ et inversement, donc que $\mathrm{rg}\mathbf{A} = \mathrm{rg}\mathbf{A}'$.

Il est d'autre part évident que les opérations élémentaires sur les lignes d'une matrice ne modifient pas le rang de la famille des lignes de cette matrice. Or, nous avons observé dans 3.3.3 que le rang de la famille des lignes d'une matrice échelonnée est égal au rang de la famille des colonnes, c'est-à-dire au rang de cette matrice. Compte tenu de 3.3.5, nous en concluons que *le rang de n'importe quelle matrice est également le rang de la famille des lignes de cette matrice.*

Comme corollaire de cette conclusion, il apparaît que

$$\mathrm{rg}\mathbf{A} \leqslant \min(m, n), \tag{3.21}$$

où $\mathbf{A}$ est une matrice de type $m \times n$.

3.3.11 Réduction d'une matrice carrée d'ordre n et de rang n

Le rang étant conservé par les opérations élémentaires sur les lignes, la forme échelonnée simplifiée de toute matrice carrée d'ordre n et de rang n est la matrice-unité $\mathbf{I}_n$.

3.3.12 Calcul du rang d'une matrice

Le rang d'une matrice peut être déterminé par réduction à la forme échelonnée. Très souvent, cependant, ce rang peut être trouvé avant que la forme échelonnée ne soit atteinte. Voici un exemple:

$$\begin{array}{l}\begin{pmatrix} 1 & -1 & 2 & -1 \\ 3 & 1 & 4 & 3 \\ -5 & -3 & -6 & -7 \\ -2 & 2 & -4 & 2 \\ -1 & 1 & 1 & 1 \end{pmatrix}\end{array}\begin{array}{l} \\ \mathbf{L}_2 - 3\mathbf{L}_1 \\ \mathbf{L}_3 + 5\mathbf{L}_1 \\ \mathbf{L}_4 + 2\mathbf{L}_1 \\ \mathbf{L}_5 + \mathbf{L}_1 \end{array} \qquad \begin{pmatrix} 1 & -1 & 2 & -1 \\ 0 & 4 & -2 & 6 \\ 0 & -8 & 4 & -12 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix}.$$

Le rang de la famille des lignes et donc de la matrice est visiblement 3.

3.4 MÉTHODE DE RÉSOLUTION DE GAUSS

3.4.1 Opérations élémentaires sur les équations d'un système linéaire

Effectuées sur les lignes de la matrice augmentée associée à un système, les opérations de type 1, 2 et 3 sont l'expression d'opérations sur les équations de ce système. Ces opérations transforment le système en un système équivalent. En effet, échanger deux équations ou multiplier une équation par un nombre non nul ne change en rien l'ensemble des solutions. D'autre part, toute solution d'un couple d'équations de la forme

$$\begin{aligned} a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n &= b_i \\ a_{l1}x_1 + a_{l2}x_2 + \ldots + a_{ln}x_n &= b_l \end{aligned}$$

est solution du couple d'équations

$$\begin{aligned} (a_{i1} + \alpha a_{l1})x_1 + (a_{i2} + \alpha a_{l2})x_2 + \ldots + (a_{in} + \alpha a_{ln})x_n &= b_i + \alpha b_l \\ a_{l1}x_1 + a_{l2}x_2 + \ldots + a_{ln}x_n &= b_l \end{aligned}$$

et inversement, ce qui montre que l'addition à une équation d'un multiple d'une autre équation laisse l'ensemble des solutions inchangé.

3.4.2 Résolution d'un système linéaire admettant exactement une solution

Considérons un système de m équations à n inconnues, de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$. Supposons que ce système admette une solution unique. Nous allons déterminer cette solution par le procédé de réduction. D'après la proposition 3.2.7, $n = \mathrm{rg}\mathbf{A} = \mathrm{rg}(\mathbf{A} \vdots \mathbf{b})$. On notera, au passage, que n est, dans ce cas, inférieur ou égal à m, par (3.21). D'après 3.3.5, la matrice augmentée $(\mathbf{A} \vdots \mathbf{b})$ peut être réduite à la forme échelonnée par des opérations élémentaires sur les lignes. Puisqu'en vertu de 3.3.10 les rangs de $\mathbf{A}$ et de $(\mathbf{A} \vdots \mathbf{b})$ ne sont pas modifiés par les opérations de réduction, la matrice échelonnée aura la forme

$$\begin{pmatrix} a'_{11} & \cdots & & \cdots & b'_1 \\ & a'_{22} & \cdots & \cdots & b'_2 \\ & & \cdots & & \\ & 0 & & & \\ & & & a'_{nn} & b'_n \end{pmatrix},$$

où $a'_{11}, a'_{22}, \ldots, a'_{nn}$ sont les pivots. Le système associé à cette matrice augmentée s'écrit

$$\begin{aligned} a'_{11}x_1 + a'_{12}x_2 + \ldots + a'_{1n}x_n &= b'_1 \\ a'_{22}x_2 + \ldots + a'_{2n}x_n &= b'_2 \\ \cdots\cdots\cdots\cdots\cdots & \\ a'_{nn}x_n &= b'_n. \end{aligned} \tag{3.22}$$

D'après 3.4.1, ce système est équivalent au système initial. Les inconnues x_1, x_2, ..., x_n peuvent être déterminées dans l'ordre inverse, en résolvant les équations successivement de la dernière à la première, par substitution.

A titre d'exemple, nous allons résoudre le système

$$\begin{aligned} x_1 + x_2 + 2x_3 - 3x_4 &= 1 \\ 2x_1 + 3x_2 + 6x_3 - 5x_4 &= 2 \\ 3x_1 - x_2 + 2x_3 - 7x_4 &= 5 \\ x_1 + 2x_2 + 3x_3 - x_4 &= -1. \end{aligned}$$

Réduisons d'abord la matrice augmentée à la forme échelonnée:

$$\begin{pmatrix} 1 & 1 & 2 & -3 & 1 \\ 2 & 3 & 6 & -5 & 2 \\ 3 & -1 & 2 & -7 & 5 \\ 1 & 2 & 3 & -1 & -1 \end{pmatrix}\begin{array}{l} \\ \mathbf{L}_2 - 2\mathbf{L}_1 \\ \mathbf{L}_3 - 3\mathbf{L}_1 \\ \mathbf{L}_4 - \mathbf{L}_1 \end{array} \qquad \begin{pmatrix} 1 & 1 & 2 & -3 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & -4 & -4 & 2 & 2 \\ 0 & 1 & 1 & 2 & -2 \end{pmatrix}\begin{array}{l} \\ \\ \frac{1}{2}(\mathbf{L}_3 + 4\mathbf{L}_2) \\ \mathbf{L}_4 - \mathbf{L}_2 \end{array}$$

$$\begin{pmatrix} 1 & 1 & 2 & -3 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 3 & 1 \\ 0 & 0 & -1 & 1 & -2 \end{pmatrix}\begin{array}{l} \\ \\ \\ 2\mathbf{L}_4 + \mathbf{L}_3 \end{array} \qquad \begin{pmatrix} 1 & 1 & 2 & -3 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 3 & 1 \\ 0 & 0 & 0 & 5 & -3 \end{pmatrix}.$$

Le système associé à cette matrice augmentée s'écrit

$$\begin{aligned} x_1 + x_2 + 2x_3 - 3x_4 &= 1 \\ x_2 + 2x_3 + x_4 &= 0 \\ 2x_3 + 3x_4 &= 1 \\ 5x_4 &= -3. \end{aligned}$$

La dernière équation détermine x_4, puis la troisième x_3 et ainsi de suite. L'unique solution du système est

$$x_1 = -\frac{7}{5},\ x_2 = -\frac{11}{5},\ x_3 = \frac{7}{5},\ x_4 = -\frac{3}{5}.$$

3.4.3 Résolution d'un système linéaire dans le cas général

Considérons à nouveau un système de m équations à n inconnues, de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$. Si la matrice $\mathbf{A}$ est nulle, ce système n'admet de solutions qu'à la condition que $\mathbf{b}$ aussi soit nul. Sous cette condition, l'ensemble des solutions est manifestement $\mathbb{R}^n$. Si la matrice $\mathbf{A}$ est non nulle, nous transformons la matrice augmentée $(\mathbf{A} \vdots \mathbf{b})$ en une matrice échelonnée, selon ce qui a été établi dans 3.3.5. Le résultat de cette transformation est soit une matrice de la forme

$$\begin{pmatrix} a'_{1j_1} & \cdots & & & \cdots & a'_{1n} & b'_1 \\ & & a'_{2j_2} & \cdots & \cdots & a'_{2n} & b'_2 \\ & & & \cdots\cdots\cdots & & & \\ 0 & & & & a'_{rj_r} \ \cdots & a'_{rn} & b'_r \\ & & & & & & \end{pmatrix}, \qquad (3.23)$$

où $a'_{1j_1}, a'_{2j_2}, \ldots, a'_{rj_r}$ sont les pivots, soit une matrice de cette même forme, mais avec un pivot supplémentaire b'_{r+1} dans la dernière colonne. Dans le deuxième cas, le système n'admet aucune solution, car la dernière colonne n'est pas combinaison linéaire des autres colonnes. Dans le premier cas, le rang de la matrice associée au système est égal au rang de la matrice augmentée, donc le système admet au moins une solution, d'après la proposition 3.2.7. Pour trouver l'ensemble des solutions, nous écrivons d'abord le système réduit équivalent au système initial:

$$\begin{aligned} a'_{1j_1}x_{j_1} + \ldots \qquad\qquad \ldots + a'_{1n}x_n &= b'_1 \\ a'_{2j_2}x_{j_2} + \ldots \qquad \ldots + a'_{2n}x_n &= b'_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a'_{rj_r}x_{j_r} + \ldots + a'_{rn}x_n &= b'_r. \end{aligned} \qquad (3.24)$$

Nous attribuons ensuite des valeurs arbitraires $\alpha_1, \alpha_2, \ldots, \alpha_{n-r}$ aux $n - r$ inconnues d'indice j différent de $j_1, j_2, \ldots, j_r$ (s'il n'y a pas de telles inconnues, le système se réduit au système (3.22)). En groupant alors les termes contenant les valeurs α_j

dans le second membre, nous obtenons un système de r équations à r inconnues (appelées *inconnues principales*) que nous résolvons comme il est indiqué dans 3.4.2. La solution s'exprimera en fonction des valeurs α_j, ces valeurs jouant le rôle de paramètres.

Résolvons, par exemple, le système

$$\begin{aligned} 5x_1 + 10x_2 + 11x_3 - x_4 + x_5 - 41x_6 &= 0 \\ -3x_1 - 6x_2 - 7x_3 + x_4 - x_5 + 27x_6 &= 0 \\ 6x_1 + 12x_2 + 14x_3 - 2x_4 + 2x_5 - 52x_6 &= 6 \\ 2x_1 + 4x_2 + 5x_3 - x_4 + x_5 - 19x_6 &= 3. \end{aligned} \qquad (3.25)$$

La matrice augmentée associée à ce système est la matrice (3.18). Par des opérations élémentaires sur les lignes, cette matrice a été réduite à la forme échelonnée (3.19). Le système associé à la matrice sous cette forme s'écrit

$$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 - x_5 - 3x_6 &= -6 \\ -x_3 + x_4 - x_5 + 5x_6 &= -3 \\ 2x_6 &= 6. \end{aligned} \qquad (3.26)$$

Posons

$$x_2 = \alpha_1,\ x_4 = \alpha_2,\ x_5 = \alpha_3, \qquad (3.27)$$

où α_1, α_2 et α_3 sont des nombres arbitraires. Le système

$$\begin{aligned} x_1 + x_3 - 3x_6 &= -6 - 2\alpha_1 - \alpha_2 + \alpha_3 \\ -x_3 + 5x_6 &= -3 - \alpha_2 + \alpha_3 \\ 2x_6 &= 6 \end{aligned}$$

peut alors être résolu, par substitution, de la dernière à la première équation. La solution générale est

$$\begin{aligned} x_1 &= -15 - 2\alpha_1 - 2\alpha_2 + 2\alpha_3 \\ x_2 &= \alpha_1 \\ x_3 &= 18 + \alpha_2 - \alpha_3 \\ x_4 &= \alpha_2 \\ x_5 &= \alpha_3 \\ x_6 &= 3. \end{aligned} \qquad (3.28)$$

3.4.4 Utilisation de la forme échelonnée simplifiée

Lorsque le système à résoudre admet plus d'une solution, il est souvent plus commode de poursuivre la réduction jusqu'à la forme échelonnée simplifiée. On peut alors résoudre le système réduit, comme dans 3.4.3, en attribuant des valeurs arbitraires aux inconnues dont l'indice n'est pas l'indice de colonne d'un pivot et en exprimant les autres inconnues en fonction des valeurs attribuées.

Par exemple, le système (3.25) est équivalent au système associé à la matrice augmentée (3.20) (obtenue par réduction de la matrice (3.18) à la forme échelonnée simplifiée):

$$\begin{aligned} x_1 + 2x_2 \quad + 2x_4 - 2x_5 \quad &= -15 \\ x_3 - x_4 + x_5 \quad &= 18 \\ x_6 &= 3. \end{aligned}$$

On résout ce système en posant $x_2 = \alpha_1, x_4 = \alpha_2, x_5 = \alpha_3$ et en écrivant, sans plus aucun calcul, x_1, x_3, x_6 en fonction de $\alpha_1, \alpha_2, \alpha_3$. Le résultat est la solution (3.28).

3.4.5 Résolution simultanée de plusieurs systèmes linéaires

Pour résoudre plusieurs systèmes associés à une même matrice $\mathbf{A} = (a_{ij})$ et dont les seconds membres sont $\mathbf{b} = (b_i)$, $\mathbf{c} = (c_i)$, $\mathbf{d} = (d_i)$, ..., il y a avantage à effectuer les opérations de réduction une seule fois sur la matrice $\mathbf{A}$ augmentée des colonnes $\mathbf{b}, \mathbf{c}, \mathbf{d}, \dots$.

Par exemple, pour résoudre les deux systèmes

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 1 \\ 3x_1 + 2x_2 + 3x_3 &= 4 \\ 5x_1 + 4x_2 + 7x_3 &= 6, \end{aligned} \qquad \begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ 3x_1 + 2x_2 + 3x_3 &= -3 \\ 5x_1 + 4x_2 + 7x_3 &= 3, \end{aligned}$$

on réduit la matrice associée, augmentée des deux seconds membres, à la forme échelonnée simplifiée:

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 3 \\ 3 & 2 & 3 & 4 & -3 \\ 5 & 4 & 7 & 6 & 3 \end{bmatrix} \begin{array}{l} \\ -\mathbf{L}_2 + 3\mathbf{L}_1 \\ \mathbf{L}_3 - 5\mathbf{L}_1 \end{array}$$

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & -1 & 12 \\ 0 & -1 & -3 & 1 & -12 \end{bmatrix} \begin{array}{l} \mathbf{L}_1 - \mathbf{L}_2 \\ \\ \mathbf{L}_3 + \mathbf{L}_2 \end{array}$$

$$\begin{bmatrix} 1 & 0 & -1 & 2 & -9 \\ 0 & 1 & 3 & -1 & 12 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Les deux systèmes réduits s'écrivent alors

$$\begin{aligned} x_1 \quad - x_3 &= 2 \\ x_2 + 3x_3 &= -1, \end{aligned} \qquad \begin{aligned} x_1 \quad - x_3 &= -9 \\ x_2 + 3x_3 &= 12. \end{aligned}$$

On obtient les solutions en attribuant une valeur arbitraire α à l'inconnue x_3 et en exprimant les deux autres inconnues en fonction de α:

$$\begin{aligned} x_1 &= 2 + \alpha \\ x_2 &= -1 - 3\alpha \\ x_3 &= \alpha, \end{aligned} \qquad \begin{aligned} x_1 &= -9 + \alpha \\ x_2 &= 12 - 3\alpha \\ x_3 &= \alpha. \end{aligned}$$

3.4.6 Choix du pivot

Lorsqu'on résout un système par réduction à la forme échelonnée en calculant avec un nombre limité de chiffres significatifs, on amoindrit les erreurs d'arrondi en choisissant, parmi les coefficients pouvant jouer le rôle de pivot, le plus grand en valeur absolue. Le choix d'un pivot petit en valeur absolue diminue la précision des calculs. Les erreurs proviennent, en effet, de l'addition à certains coefficients de multiples d'autres coefficients divisés par un pivot. Ces erreurs seront d'autant plus petites que le pivot est grand en valeur absolue.

3.4.7 Nombre d'opérations

Le nombre d'opérations nécessaires pour résoudre un système de n équations à n inconnues par réduction de la matrice augmentée associée à la forme échelonnée est, pour n grand, de l'ordre de n^3. La preuve de cette assertion est laissée comme exercice.

3.5 STRUCTURE ET DIMENSION DE L'ENSEMBLE DES SOLUTIONS

3.5.1 Introduction

Pour étudier l'ensemble des solutions d'un système, il est utile de considérer toute solution comme un vecteur-colonne de $\mathbb{R}^n$, où n est le nombre d'inconnues de ce système. L'ensemble des solutions est alors un sous-ensemble de $\mathbb{R}^n$ dont nous allons examiner la structure.

3.5.2 Systèmes linéaires homogènes

Considérons un système homogène de m équations à n inconnues et de matrice associée $\mathbf{A} = (a_{ij})$. Il est clair que $\mathbf{0}$ est une solution (appelée *solution nulle* ou *triviale*). En outre, toute combinaison linéaire $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 = (\alpha_1 x_j^1 + \alpha_2 x_j^2)$ de solutions $\mathbf{x}_1 = (x_j^1)$ et $\mathbf{x}_2 = (x_j^2)$ est également une solution, car

$$\sum_{j=1}^{n} a_{ij}(\alpha_1 x_j^1 + \alpha_2 x_j^2) = \alpha_1 \sum_{j=1}^{n} a_{ij} x_j^1 + \alpha_2 \sum_{j=1}^{n} a_{ij} x_j^2 = 0, \quad i = 1, 2, \dots, m.$$

Cela montre que l'ensemble des solutions est un sous-espace vectoriel de $\mathbb{R}^n$. Si $\mathbf{A}$ est la matrice nulle, il est évident que ce sous-espace est $\mathbb{R}^n$ lui-même. Ecartons ce cas et réduisons $\mathbf{A}$ à la forme échelonnée simplifiée. Supposons, pour alléger l'exposition, que la matrice réduite soit de la forme particulière

$$\begin{pmatrix} 1 & a'_{12} & 0 & a'_{14} & 0 & a'_{16} & a'_{17} \\ 0 & 0 & 1 & a'_{24} & 0 & a'_{26} & a'_{27} \\ 0 & 0 & 0 & 0 & 1 & a'_{36} & a'_{37} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.29}$$

En procédant comme il est indiqué dans 3.4.4, c'est-à-dire en attribuant aux inconnues x_2, x_4, x_6, x_7 des valeurs arbitraires α_1, α_2, α_3, α_4, nous voyons que la solution générale du système s'exprime par

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \alpha_1 \begin{pmatrix} -a'_{12} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -a'_{14} \\ 0 \\ -a'_{24} \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} -a'_{16} \\ 0 \\ -a'_{26} \\ 0 \\ -a'_{36} \\ 1 \\ 0 \end{pmatrix} + \alpha_4 \begin{pmatrix} -a'_{17} \\ 0 \\ -a'_{27} \\ 0 \\ -a'_{37} \\ 0 \\ 1 \end{pmatrix}. \tag{3.30}$$

Désignons les quatre vecteurs-colonnes du second membre de (3.30) par $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ et $\mathbf{x}_4$. Ces vecteurs sont des solutions du système ($\mathbf{x}_i$ s'obtient en posant $\alpha_i = 1$ et $\alpha_j = 0$ pour $j \neq i$) et, de plus, ils sont linéairement indépendants. En outre, d'après (3.30), toute solution du système est combinaison linéaire de ces quatre vecteurs. Ils forment donc une base du sous-espace vectoriel des solutions, ce sous-espace étant par conséquent de dimension 4. On remarquera que 4 est égal au nombre d'inconnues moins le rang de A. Le raisonnement dans le cas général est le même et entraîne la proposition 3.5.3.

3.5.3 Proposition

Les solutions d'un système linéaire homogène à n inconnues forment un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n - r$, où r est le rang de la matrice associée au système.

Le corollaire suivant dérive également de la proposition 3.2.7.

3.5.4 Corollaire

Pour qu'un système linéaire homogène à n inconnues n'admette que la solution nulle, il faut et il suffit que $r = n$, où r est le rang de la matrice associée au système.

En raison de (3.21), un système homogène dont le nombre d'équations est inférieur au nombre d'inconnues admet donc toujours des solutions non nulles.

3.5.5 Systèmes linéaires dont le second membre n'est pas nul

Considérons un système de m équations à n inconnues, de matrice associée $\mathrm{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$. Supposons que ce système admette au moins une solution $\mathbf{x}_0 = (x_j^0)$. Si $\mathbf{y} = (y_j)$ est aussi une solution, alors

$$\sum_{j=1}^{n} a_{ij}(y_j - x_j^0) = \sum_{j=1}^{n} a_{ij}y_j - \sum_{j=1}^{n} a_{ij}x_j^0 = b_i - b_i = 0, \quad i = 1, 2, \dots, m,$$

ce qui montre que $\mathbf{y} - \mathbf{x}_0 = (y_j - x_j^0)$ est une solution du système homogène associé et donc que toute solution du système s'obtient en additionnant à $\mathbf{x}_0$ une solution du système homogène associé. Compte tenu de la proposition 3.5.3, la solution générale s'écrit donc sous la forme

$$\mathbf{x} = \mathbf{x}_0 + \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \dots + \alpha_{n-r}\mathbf{x}_{n-r}, \tag{3.31}$$

où $\alpha_1, \alpha_2, \dots, \alpha_{n-r}$ sont des nombres arbitraires et $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-r}$ forment une base du sous-espace vectoriel des solutions du système homogène associé (si $r = n$, la solution $\mathbf{x}_0$ est unique). Il en résulte que l'ensemble des solutions du système est un sous-espace affine de $\mathbb{R}^n$, plus exactement le sous-espace affine

$$\mathbf{x}_0 + S, \tag{3.32}$$

où S désigne le sous-espace vectoriel des solutions du système homogène associé (cf. exemples (3) et (4) de 1.9.7).

Les résultats obtenus sont résumés dans les paragraphes 3.5.6 à 3.5.9.

3.5.6 Proposition

Toute solution d'un système linéaire admettant au moins une solution $\mathbf{x}_0$ s'obtient en additionnant à $\mathbf{x}_0$ une solution du système homogène associé.

3.5.7 Proposition

L'ensemble des solutions d'un système linéaire à n inconnues admettant au moins une solution est un sous-espace affine de $\mathbb{R}^n$ de dimension $n - r$, où r est le rang de la matrice associée au système.

D'après cette proposition, un système linéaire dont le nombre d'équations m est inférieur au nombre d'inconnues n n'admet en aucun cas une seule solution, car $n - r > m - r \geqslant 0$.

3.5.8 Corollaire

Pour qu'un système linéaire admette une solution unique, il faut et il suffit qu'il admette au moins une solution et que le système homogène associé n'ait que la solution nulle.

Si le nombre d'équations m, le nombre d'inconnues n et le rang r sont égaux, le rang de la matrice augmentée associée au système ne peut être supérieur à r. Dans ce cas, la proposition suivante peut être considérée comme corollaire de la proposition 3.2.7.

3.5.9 Proposition

Un système linéaire de n équations à n inconnues admet une solution unique si et seulement si le rang de la matrice associée au système est n.

3.6 EXERCICES

3.6.1 A quelle condition doit satisfaire le coefficient α, pour que le système

$$\begin{aligned} x_2 + x_3 &= b_1 \\ x_1 \quad\ + x_3 &= b_2 \\ x_1 + x_2 + \alpha x_3 &= b_3 \end{aligned}$$

admette au moins une solution quel que soit le choix de b_1, b_2 et b_3? Cette condition étant remplie, ce système a-t-il plus d'une solution?

3.6.2 Dans chacun des cas suivants, dire, sans le résoudre, si le système n'a aucune solution, une seule solution, plus d'une solution:

(a) $$\begin{aligned} 2x_1 - 2x_2 + 4x_3 &= 1 \\ -3x_1 + 3x_2 - 6x_3 &= 2 \\ x_1 - x_2 + 2x_3 &= 1. \end{aligned}$$

(b) $$\begin{aligned} 5x_1 - 3x_2 - x_3 + 4x_4 &= 6 \\ -x_1 + 2x_2 + x_3 - x_4 &= -2 \\ 2x_2 - 2x_3 + 3x_4 &= 3. \end{aligned}$$

(c) $$\begin{aligned} 2x_1 + x_2 + 5x_3 &= 3 \\ x_1 - 2x_2 - 2x_3 &= -1 \\ -3x_1 + 2x_2 + x_3 &= -1 \\ 3x_1 - 3x_2 - 2x_3 &= 0. \end{aligned}$$

3.6.3 Déterminer le polynôme p de degré 3 qui satisfait à l'équation différentielle $\ddot{p}(t) + \dot{p}(t) - p(t) + t - t^3 \equiv 0$, où le point désigne la dérivation par rapport à t.

3.6.4 Réduire les matrices suivantes à la forme échelonnée:

$$\begin{bmatrix} 3 & 2 & 5 & -1 & 2 \\ 4 & -1 & 2 & 3 & -3 \\ -2 & 3 & -1 & 1 & 4 \\ 5 & -2 & 2 & 1 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 2 & -1 & -1 & 4 & -1 & 2 \\ -3 & 2 & 1 & -3 & 2 & -3 \\ 4 & 3 & -4 & 2 & 0 & 5 \\ 5 & 1 & 0 & -3 & 4 & -2 \end{bmatrix}.$$

3.6.5 Décrire l'ensemble des triplets (a, b, c) rendant le rang de la matrice

$$\begin{bmatrix} 1 & 1 & a & b \\ -1 & 2 & b & c \\ 2 & -1 & c & a \end{bmatrix}$$

égal à 2.

3.6.6 Soit $a_1, a_2, ..., a_n$ des nombres distincts et k un entier positif inférieur à n. Montrer que le rang de la matrice

$$\begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^k \\ 1 & a_2 & a_2^2 & \dots & a_2^k \\ & & \dots\dots\dots & & \\ 1 & a_n & a_n^2 & \dots & a_n^k \end{bmatrix}$$

est égal à $k+1$.

3.6.7 A quelle condition doivent satisfaire b_1, b_2 et b_3 pour que le système

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= b_1 \\ 2x_1 + 6x_2 - 11x_3 &= b_2 \\ x_1 - 2x_2 + 7x_3 &= b_3 \end{aligned}$$

admette au moins une solution?

3.6.8 Dans un espace affine de dimension 4 muni d'un repère, on considère les deux plans d'équations paramétriques respectives

$$\begin{aligned} x_1 &= 1 + \alpha_1 + 2\alpha_2 \\ x_2 &= -1 - \alpha_1 \\ x_3 &= 2 + 3\alpha_1 - \alpha_2 \\ x_4 &= 1 + \alpha_1 + \alpha_2, \end{aligned} \qquad \begin{aligned} x_1 &= 2 + \alpha_1' + 2\alpha_2' \\ x_2 &= -2 \quad\ + 2\alpha_2' \\ x_3 &= 5 + \alpha_1' - 5\alpha_2' \\ x_4 &= 2 + 2\alpha_1' + 3\alpha_2'. \end{aligned}$$

Trouver l'intersection de ces deux plans.

3.6.9 Réduire la matrice augmentée associée au système

$$\begin{aligned} 3x_1 - 3x_2 - x_3 + 2x_4 - 9x_5 &= 13 \\ x_1 - x_2 + 2x_3 - x_4 - 6x_5 &= -6 \\ x_1 - x_2 + x_3 + x_4 - 6x_5 &= 1 \\ -x_1 + x_2 - x_3 - 2x_4 + 7x_5 &= -3 \end{aligned}$$

à la forme échelonnée, puis en tirer la solution générale du système.

3.6.10 Résoudre le système

$$\begin{aligned} ax_1 + bx_2 + cx_3 &= 1 \\ a^2x_1 + b^2x_2 + c^2x_3 &= 1 \\ a^3x_1 + b^3x_2 + c^3x_3 &= 1. \end{aligned}$$

où a, b et c sont des nombres distincts et non nuls.

3.6.11 Par réduction simultanée, résoudre les deux systèmes associés à la matrice

$$\begin{bmatrix} 1 & 5 & 2 \\ -1 & 2 & 1 \\ 5 & 4 & 1 \end{bmatrix},$$

dont les seconds membres respectifs sont $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ et $\begin{bmatrix} 2 \\ 3 \\ -5 \end{bmatrix}$.

3.6.12 Résoudre le système

$$\begin{aligned}
x_1 &= \quad\; bx_2 && + 1 \\
x_2 &= ax_1 + bx_3 && + 1 \\
x_3 &= \quad\; ax_2 + bx_4 && + 1 \\
&\quad \cdots\cdots\cdots\cdots\cdots\cdots \\
x_{n-1} &= \quad\; ax_{n-2} + bx_n && + 1 \\
x_n &= \quad\; ax_{n-1} && + 1,
\end{aligned}$$

où a et b sont des nombres distincts non nuls dont la somme est 1. (Poser $x_0 = x_{n+1} = 0$, introduire de nouvelles inconnues $y_j = x_j - x_{j-1}$, $j = 1, 2, \ldots, n+1$, et résoudre récursivement.)

Chapitre 4

Algèbre matricielle

4.1 OPÉRATIONS SUR LES MATRICES

4.1.1 Introduction

Définies dans le chapitre précédent en vue de faciliter l'étude des systèmes linéaires, les matrices deviendront, dès maintenant, des objets pouvant se lier les uns aux autres, en accord avec des règles que nous établirons.

4.1.2 Addition de matrices

Soit $\mathbf{A} = (a_{ij})$ et $\mathbf{B} = (b_{ij})$ des matrices de type $m \times n$. On appelle *somme* de $\mathbf{A}$ et $\mathbf{B}$, et on note $\mathbf{A} + \mathbf{B}$, la matrice $\mathbf{C} = (c_{ij})$ de type $m \times n$ dont les termes sont définis par la relation

$$c_{ij} = a_{ij} + b_{ij}.$$

L'opération qui associe à tout couple de matrices du même type leur somme est appelée *addition matricielle*. Exemple:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{bmatrix}.$$

De toute évidence, l'opération d'addition matricielle est associative et commutative:

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}), \quad \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.$$

L'élément neutre est la matrice nulle et la matrice opposée de $\mathbf{A} = (a_{ij})$ est la matrice $-\mathbf{A} = (-a_{ij})$.

Si $\mathbf{A}$ et $\mathbf{B}$ sont des matrices du même type, la matrice $\mathbf{A} + (-\mathbf{B})$ est notée plus simplement $\mathbf{A} - \mathbf{B}$ et appelée *différence* de $\mathbf{A}$ et $\mathbf{B}$. L'opération qui associe à tout couple de matrices du même type leur différence est appelée *soustraction matricielle*.

L'addition matricielle s'étend de manière évidente au cas d'une famille finie quelconque de matrices du même type.

4.1.3 Multiplication d'une matrice par un scalaire

On appelle *produit du nombre* α *par la matrice* **A** la matrice, notée $\alpha\mathbf{A}$, dont les termes sont ceux de **A** multipliés par α. L'opération consistant à effectuer le produit d'un nombre par une matrice est appelée *multiplication par un scalaire.* Exemple:

$$\alpha\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \alpha a_{13} \\ \alpha a_{21} & \alpha a_{22} & \alpha a_{23} \end{pmatrix}.$$

Les propriétés suivantes découlent immédiatement des définitions:

(a) $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$.
(b) $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$.
(c) $\alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A}$.
(d) $(-1)\mathbf{A} = -\mathbf{A},\ 0\mathbf{A} = \mathbf{O},\ \alpha\mathbf{O} = \mathbf{O}$.

4.1.4 Espaces vectoriels de matrices

Muni des opérations d'addition et de multiplication par un scalaire, l'ensemble des matrices de type $m \times n$ devient un espace vectoriel de dimension mn. Les matrices dont un des termes est égal à 1 et les autres sont nuls forment une base de cet espace appelée *base canonique.*

4.1.5 Multiplication de matrices

Soit $\mathbf{A} = (a_{ij})$ une matrice de type $m \times n$ et $\mathbf{B} = (b_{jk})$ une matrice de type $n \times p$. On appelle *produit* de **A** et **B**, et on note **AB**, la matrice $\mathbf{C} = (c_{ik})$ de type $m \times p$ dont les termes sont définis par la relation

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}. \qquad (4.1)$$

L'opération qui associe à tout couple de matrices leur produit (lorsque celui-ci est défini) est appelée *multiplication matricielle.* Exemples:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} & a_{11}b_{13}+a_{12}b_{23} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} & a_{21}b_{13}+a_{22}b_{23} \end{pmatrix},$$

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 + a_2b_2 + \dots + a_nb_n \end{pmatrix},$$

$$\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}\begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_n \\ a_2b_1 & a_2b_2 & \dots & a_2b_n \\ \dots & \dots & \dots & \dots \\ a_nb_1 & a_nb_2 & \dots & a_nb_n \end{pmatrix}.$$

On retiendra que le produit **AB** n'est défini que si le nombre de colonnes de **A** est égal au nombre de lignes de **B**. Schématiquement, cela peut s'exprimer ainsi:

$${}_m\!\updownarrow\overset{\overset{p}{\leftrightarrow}}{\mathbf{C}} = {}_m\!\updownarrow\overset{\overset{n}{\leftrightarrow}}{\mathbf{A}}\ {}_n\!\updownarrow\overset{\overset{p}{\leftrightarrow}}{\mathbf{B}},$$

ou encore

$$m \times p = (m \times n)\cdot(n \times p).$$

En particulier, la multiplication est toujours possible entre matrices carrées du même ordre.

Le terme d'indices i, k du produit **AB** est, selon (4.1), le produit scalaire de la i-ième ligne de **A** par la k-ième colonne de **B**, toutes deux considérées comme des vecteurs de $\mathbb{R}^n$.

La multiplication matricielle n'est pas une opération commutative. En fait, le produit **BA** n'est pas nécessairement défini lorsque **AB** l'est. D'autre part, les produits **AB** et **BA** peuvent être définis et ne pas être du même type. Mais même lorsqu'ils sont du même type, ils sont généralement différents, comme le montre l'exemple simple suivant:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

En revanche, la multiplication matricielle est une opération associative et distributive. Plus précisément, chacune des égalités suivantes est vraie, à la seule condition que l'un de ses deux membres soit défini (l'autre l'est alors également):

(a) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ (associativité).
(b) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
(c) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ (distributivité).

Il est en outre évident qu'à la même condition,

(d) $(\alpha\mathbf{A})(\beta\mathbf{B}) = (\alpha\beta)\mathbf{AB}$.

A titre d'exemple, démontrons (a). Soit $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{jk})$ et $\mathbf{C} = (c_{kl})$ des matrices respectivement de type $m \times n$, $n \times p$ et $p \times q$. Le terme d'indices i, l de $\mathbf{A}(\mathbf{BC})$ est

$$\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{p} b_{jk}c_{kl}\right)$$

et le terme de mêmes indices de $(\mathbf{AB})\mathbf{C}$ est

$$\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)c_{kl}.$$

S'agissant de sommes finies, l'ordre des sommations sur j et sur k peut être inversé, ce qui montre l'égalité des deux expressions.

4.1.6 Produits de plusieurs facteurs et puissances

On définit un produit matriciel de plusieurs facteurs par une succession de produits de deux facteurs. Grâce à l'associativité, un tel produit peut s'écrire sans parenthèses. Par exemple, $\mathbf{ABCD}$ désigne la succession de produits $((\mathbf{AB})\mathbf{C})\mathbf{D}$.

Si $\mathbf{A}$ est une matrice carrée, le produit $\mathbf{AA} \dots \mathbf{A}$ (k facteurs) est désigné par $\mathbf{A}^k$ et appelé *puissance k-ième* de $\mathbf{A}$. Par convention, $\mathbf{A}^0 = \mathbf{I}$. Il est évident que

$$\mathbf{A}^k\mathbf{A}^l = \mathbf{A}^{k+l} \quad \text{et} \quad (\mathbf{A}^k)^l = \mathbf{A}^{kl}, \tag{4.2}$$

où k et l sont des entiers non négatifs.

4.1.7 Exemples et remarques

(1) Voici deux exemples de multiplication par la matrice-unité $\mathbf{I}_3$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

D'une manière générale, on constate aisément que la matrice-unité joue le rôle d'élément neutre de la multiplication matricielle, autrement dit que

$$\mathbf{AI}_n = \mathbf{A} \quad \text{et} \quad \mathbf{I}_m\mathbf{A} = \mathbf{A},$$

pour toute matrice $\mathbf{A}$ de type $m \times n$. Grâce à la propriété (d) de 4.1.5, il s'ensuit que

$$\mathbf{A}(\alpha\mathbf{I}_n) = \alpha\mathbf{A} \quad \text{et} \quad (\alpha\mathbf{I}_m)\mathbf{A} = \alpha\mathbf{A}.$$

(2) On dit que deux matrices carrées du même ordre $\mathbf{A}$ et $\mathbf{B}$ *commutent*, ou que l'une des deux *commute avec* l'autre, si $\mathbf{AB} = \mathbf{BA}$. Nous venons de voir que les matrices $\alpha\mathbf{I}_n$ commutent avec toutes les matrices carrées d'ordre n. Ces matrices sont d'ailleurs les seules qui jouissent de cette propriété (cf. exercice 4.7.5). A noter que si $\mathbf{A}$ et $\mathbf{B}$ commutent,

$$(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2.$$

Plus généralement, sous la même hypothèse, il est facile de constater, par récurrence sur k, que

$$(\mathbf{A} + \mathbf{B})^k = \sum_{l=0}^{k}\binom{k}{l}\mathbf{A}^l\mathbf{B}^{k-l} \quad \textit{(formule du binôme)}, \tag{4.3}$$

où k désigne un entier positif.

(3) Il est évident que $\mathbf{AB} = \mathbf{O}$ si $\mathbf{A} = \mathbf{O}$ ou $\mathbf{B} = \mathbf{O}$. Mais le produit $\mathbf{AB}$ peut être nul sans que $\mathbf{A}$ ou $\mathbf{B}$ soient des matrices nulles, comme le montre l'exemple simple suivant:

$$\begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix}\begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Il est donc aussi possible que

$$\mathbf{AB} = \mathbf{AC},$$

sans que $\mathbf{A} = \mathbf{O}$ ou que $\mathbf{B} = \mathbf{C}$.

(4) En posant $a = d = 1$ et $b = c = 0$ dans l'exemple ci-dessus, nous voyons que le carré d'une matrice non nulle peut être nul. A l'aide de la multiplication par blocs, nous verrons que la puissance n-ième de toute matrice carrée de la forme

$$\mathbf{A}_n = \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ & & \dots\dots & & \\ 0 & 0 & 0 & \dots & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{4.4}$$

est la matrice nulle.

D'une manière générale, on appelle matrice *nilpotente* toute matrice carrée $\mathbf{A}$ telle que $\mathbf{A}^k = \mathbf{O}$ pour un entier positif k.

(5) L'opération de multiplication matricielle permet d'exprimer tout système linéaire sous la forme d'une équation matricielle. En effet, résoudre un système de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$ revient à chercher les matrices-colonnes $\mathbf{x} = (x_j)$ telles que

$$\mathbf{Ax} = \mathbf{b},$$

ou encore,

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots\dots\dots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

(6) Soit $\mathbf{A}$ une matrice de type $m \times n$. Si $\mathbf{Ax} = \mathbf{0}$ pour tout vecteur-colonne $\mathbf{x}$ de $\mathbb{R}^n$, alors $\mathbf{A}$ est la matrice nulle. Cela découle immédiatement de la proposition 3.5.3, mais se démontre également en observant que $\mathbf{Ax}$ est la j-ième colonne de $\mathbf{A}$ si $\mathbf{x}$ est le j-ième vecteur-colonne de la base canonique de $\mathbb{R}^n$.

L'assertion suivante en est un corollaire:

Soit $\mathbf{A}$ et $\mathbf{B}$ des matrices de type $m \times n$. Si $\mathbf{Ax} = \mathbf{Bx}$ pour tout vecteur-colonne $\mathbf{x}$ de $\mathbb{R}^n$, alors $\mathbf{A} = \mathbf{B}$.

(7) On appelle *trace* d'une matrice carrée $\mathbf{A}$, et on note $\operatorname{tr}\mathbf{A}$, la somme des termes diagonaux de $\mathbf{A}$.

Si $\mathbf{A} = (a_{ij})$ et $\mathbf{B} = (b_{ij})$ sont deux matrices carrées d'ordre n,

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}). \tag{4.5}$$

En effet,

$$\operatorname{tr}(\mathbf{AB}) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{ji}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} b_{ji}a_{ij}\right) = \operatorname{tr}(\mathbf{BA}).$$

4.1.8 Multiplication matricielle par blocs

On dit qu'une matrice $\mathbf{A}$ est partagée en sous-matrices si elle est de la forme

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1r} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2r} \\ & \cdots\cdots & & \\ \mathbf{A}_{q1} & \mathbf{A}_{q2} & \cdots & \mathbf{A}_{qr} \end{bmatrix}, \tag{4.6}$$

où $\mathbf{A}_{ij}$ est, pour $i = 1, 2, ..., q$ et $j = 1, 2, ..., r$, une matrice dont le nombre de lignes est indépendant de j et le nombre de colonnes indépendant de i.

On peut multiplier des matrices partagées en sous-matrices, en opérant comme si chaque sous-matrice était un terme ordinaire d'une matrice. La seule condition que doivent remplir les partitions est que les produits matriciels entrant en jeu soient définis. On appelle ce mode de multiplication *multiplication matricielle par blocs*.

Voici un exemple:

$$\begin{array}{c} \\ m \\ m' \end{array}\begin{array}{c} \begin{array}{cc} n & n' \end{array} \\ \left[\begin{array}{c:c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hdashline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right] \end{array}
\begin{array}{c} \\ n \\ n' \end{array}\begin{array}{c} \begin{array}{ccc} p & p' & p'' \end{array} \\ \left[\begin{array}{c:c:c} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \hdashline \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{array}\right] \end{array}
=
\begin{array}{c} \\ m \\ m' \end{array}\begin{array}{c} \begin{array}{ccc} p & p' & p'' \end{array} \\ \left[\begin{array}{c:c:c} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} & \mathbf{A}_{11}\mathbf{B}_{13} + \mathbf{A}_{12}\mathbf{B}_{23} \\ \hdashline \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} & \mathbf{A}_{21}\mathbf{B}_{13} + \mathbf{A}_{22}\mathbf{B}_{23} \end{array}\right] \end{array}.$$

D'une manière générale, si $\mathbf{A}$ est la matrice (4.6) et

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1s} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \cdots & \mathbf{B}_{2s} \\ & \cdots\cdots & & \\ \mathbf{B}_{r1} & \mathbf{B}_{r2} & \cdots & \mathbf{B}_{rs} \end{bmatrix},$$

sous réserve que les produits $\mathbf{A}_{ij}\mathbf{B}_{jk}$ soient définis pour tout choix des indices i, j, k,

$$\mathbf{AB} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1s} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{2s} \\ & \cdots\cdots & & \\ \mathbf{C}_{q1} & \mathbf{C}_{q2} & \cdots & \mathbf{C}_{qs} \end{bmatrix},$$

où

$$\mathbf{C}_{ik} = \mathbf{A}_{i1}\mathbf{B}_{1k} + \mathbf{A}_{i2}\mathbf{B}_{2k} + \dots + \mathbf{A}_{ir}\mathbf{B}_{rk}. \tag{4.7}$$

En vue d'applications ultérieures, nous allons utiliser la multiplication par blocs pour calculer le produit $\mathbf{AB}$, sous l'hypothèse que $\mathbf{B}$ est partagée en vecteurs-colonnes. Il vient:

$$\mathbf{A}\left[\begin{array}{c:c:c:c} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{array}\right] = \left[\begin{array}{c:c:c:c} \mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_n \end{array}\right]. \tag{4.8}$$

Comme deuxième application, nous nous proposons de démontrer, par récurrence sur n, que la puissance n-ième de la matrice (4.4) est la matrice nulle. Pour $n = 1$ il n'y a rien à démontrer. Supposons que $\mathbf{A}_{n-1}^{n-1} = \mathbf{O}$ et partageons $\mathbf{A}_n$ de la manière suivante:

$$\mathbf{A}_n = \left[\begin{array}{c:c} & a_{1n} \\ \mathbf{A}_{n-1} & \vdots \\ & a_{n-1,n} \\ \hdashline 0\ 0\ \ldots\ 0 & 0 \end{array}\right].$$

En multipliant par blocs, nous voyons aisément que

$$\mathbf{A}_n^2 = \left[\begin{array}{c:c} & \times \\ \mathbf{A}_{n-1}^2 & \vdots \\ & \times \\ \hdashline 0\ 0\ \ldots\ 0 & 0 \end{array}\right], \quad \ldots, \quad \mathbf{A}_n^{n-1} = \left[\begin{array}{c:c} & \times \\ \mathbf{A}_{n-1}^{n-1} & \vdots \\ & \times \\ \hdashline 0\ 0\ \ldots\ 0 & 0 \end{array}\right],$$

où les croix occupent la place de termes numériques qu'il n'est pas nécessaire de spécifier. Mais alors

$$\mathbf{A}_n^n = \mathbf{A}_n^{n-1}\mathbf{A}_n = \left(\begin{array}{ccccc|c} & & & & & \times \\ & & & & & \cdot \\ & & \mathbf{O} & & & \cdot \\ & & & & & \cdot \\ & & & & & \times \\ \hline 0 & 0 & . & . & . \; 0 & 0 \end{array}\right)\left(\begin{array}{ccccc|c} & & & & & a_{1n} \\ & & & & & \cdot \\ & & \mathbf{A}_{n-1} & & & \cdot \\ & & & & & \cdot \\ & & & & & a_{n-1,n} \\ \hline 0 & 0 & . & . & . \; 0 & 0 \end{array}\right) = \mathbf{O},$$

ce qui achève la démonstration.

4.1.9 Transposition d'une matrice

On appelle *transposée* d'une matrice $\mathbf{A} = (a_{ij})$, et on note ${}^t\mathbf{A} = (a'_{ij})$ la matrice dont les lignes sont les colonnes de $\mathbf{A}$ (et, par conséquent, les colonnes sont les lignes de $\mathbf{A}$). Entre les termes de ${}^t\mathbf{A}$ et ceux de $\mathbf{A}$ il y a donc la relation

$$a'_{ij} = a_{ji}. \tag{4.9}$$

En outre, ${}^t\mathbf{A}$ est de type $n \times m$ si $\mathbf{A}$ est de type $m \times n$.

Voici un exemple de transposition:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \qquad {}^t\mathbf{A} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix}.$$

On notera que l'opération de transposition laisse toute matrice de type 1×1 inchangée, transforme les matrices-lignes en matrices-colonnes et les matrices-colonnes en matrices-lignes.

Chacune des égalités suivantes est vraie chaque fois que l'un de ses deux membres est défini (l'autre l'est alors également):

(a) ${}^t(\mathbf{A} + \mathbf{B}) = {}^t\mathbf{A} + {}^t\mathbf{B}$.
(b) ${}^t(\alpha\mathbf{A}) = \alpha\,{}^t\mathbf{A}$.
(c) ${}^t({}^t\mathbf{A}) = \mathbf{A}$.
(d) ${}^t(\mathbf{AB}) = {}^t\mathbf{B}\,{}^t\mathbf{A}$.

La seule qui ne soit pas évidente est la dernière. En voici la preuve. Soit $\mathbf{A} = (a_{ij})$ une matrice de type $m \times n$ et $\mathbf{B} = (b_{jk})$ une matrice de type $n \times p$. Le terme d'indices k, i de ${}^t(\mathbf{AB})$ est le terme d'indices i, k de $\mathbf{AB}$, c'est-à-dire $\sum_{j=1}^{n} a_{ij}b_{jk}$. Le terme d'indices k, i de ${}^t\mathbf{B}\,{}^t\mathbf{A}$ est $\sum_{j=1}^{n} b'_{kj}a'_{ji}$. L'égalité a donc bien lieu en vertu de (4.9).

L'extension de (d) à une famille finie quelconque de matrices est immédiate:

(e) ${}^t(\mathbf{A}_1\mathbf{A}_2 \dots \mathbf{A}_k) = {}^t\mathbf{A}_k\,{}^t\mathbf{A}_{k-1} \dots {}^t\mathbf{A}_1$.

En particulier, lorsque $\mathbf{A}$ est une matrice carrée (et k un entier positif),

(f) ${}^t(\mathbf{A}^k) = ({}^t\mathbf{A})^k$.

4.1.10 Rang d'une matrice transposée

Pour toute matrice $\mathbf{A}$

$$\mathrm{rg}\,{}^t\mathbf{A} = \mathrm{rg}\mathbf{A}, \tag{4.10}$$

car la famille des lignes et celle des colonnes d'une même matrice ont le même rang (cf. 3.3.10).

4.2 MATRICES INVERSIBLES

4.2.1 Introduction

On sait que tout nombre non nul α admet un inverse unique, c'est-à-dire un nombre, noté α^{-1}, tel que $\alpha\alpha^{-1} = 1$. En est-il de même pour les matrices? La réponse à cette question est négative. Par exemple, le produit de la matrice

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

par une matrice quelconque

$$\mathbf{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

n'est jamais $\mathbf{I}_2$, car

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}.$$

Dans cette section, nous nous proposons d'étudier les matrices qui admettent un inverse et d'exposer une méthode de calcul de cet inverse.

4.2.2 Matrices inversibles

On dit qu'une matrice carrée $\mathbf{A}$ est *inversible* ou *régulière* s'il existe une matrice $\mathbf{B}$ (nécessairement carrée et du même ordre que $\mathbf{A}$) telle que

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}. \tag{4.11}$$

On notera que si $\mathbf{AB} = \mathbf{I}$ et $\mathbf{B'A} = \mathbf{I}$, alors $\mathbf{B} = \mathbf{B'}$, car $\mathbf{B'} = \mathbf{B'(AB)} = \mathbf{(B'A)B} = \mathbf{B}$. Lorsque $\mathbf{A}$ est inversible, la matrice $\mathbf{B}$ de la relation (4.11) est donc unique. Elle est appelée *matrice inverse* de $\mathbf{A}$ et notée $\mathbf{A}^{-1}$.

4.2.3 Proposition

Soit $\mathbf{A}$ *une matrice carrée. S'il existe une matrice* $\mathbf{B}$ *(nécessairement carrée et du même ordre que* $\mathbf{A}$*) telle que* $\mathbf{AB} = \mathbf{I}$ *ou* $\mathbf{BA} = \mathbf{I}$*, alors* $\mathbf{B}$ *vérifie* (4.11), *autrement dit,* $\mathbf{A}$ *est inversible et* $\mathbf{B} = \mathbf{A}^{-1}$.

DÉMONSTRATION

Il suffit de montrer que $\mathbf{AB} = \mathbf{I}$ entraîne $\mathbf{BA} = \mathbf{I}$. Soit n l'ordre de $\mathbf{A}$ et de $\mathbf{B}$. Le système homogène de matrice associée $\mathbf{B}$ n'admet que la solution triviale, car si $\mathbf{x}$ est une solution,

$$\mathbf{x} = \mathbf{ABx} = \mathbf{A}(\mathbf{Bx}) = \mathbf{A0} = \mathbf{0}.$$

Il s'ensuit, d'après le corollaire 3.5.4, que le rang de $\mathbf{B}$ est n. Soit maintenant $\mathbf{x}$ un vecteur-colonne quelconque de $\mathbb{R}^n$. En vertu de la proposition 3.5.9, le système de matrice associée $\mathbf{B}$ et de second membre $\mathbf{x}$ admet une solution (unique) $\mathbf{y}$. Par conséquent,

$$\mathbf{BAx} = \mathbf{BA}(\mathbf{By}) = \mathbf{B}(\mathbf{AB})\mathbf{y} = \mathbf{By} = \mathbf{x}.$$

Comme $\mathbf{x}$ est arbitraire, la remarque (6) de 4.1.7 nous permet de conclure que $\mathbf{BA} = \mathbf{I}$.

4.2.4 Proposition. Rang d'une matrice inversible

Une matrice carrée $\mathbf{A}$ *d'ordre* n *est inversible si et seulement si son rang est* n.

DÉMONSTRATION

Supposons que $\mathbf{A}$ soit inversible. Alors $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, donc le rang de $\mathbf{A}$ est n, d'après ce qui a été établi dans la première partie de la démonstration précédente. Réciproquement, supposons que le rang de $\mathbf{A}$ soit n. Le procédé exposé dans le paragraphe suivant démontre l'existence d'une matrice $\mathbf{B}$ telle que $\mathbf{AB} = \mathbf{I}$. D'après la proposition 4.2.3, $\mathbf{A}$ est donc inversible.

4.2.5 Calcul de la matrice inverse

Soit $\mathbf{A}$ une matrice carrée d'ordre n et de rang n. Nous nous proposons de résoudre l'équation matricielle

$$\mathbf{AX} = \mathbf{I}, \tag{4.12}$$

où $\mathbf{X}$ est une matrice carrée inconnue. En vertu de (4.8), cette équation s'exprime, de manière équivalente, par les n équations matricielles

$$\mathbf{Ax}_1 = \mathbf{e}_1,\ \mathbf{Ax}_2 = \mathbf{e}_2,\ \dots,\ \mathbf{Ax}_n = \mathbf{e}_n, \tag{4.13}$$

où $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ sont les colonnes de $\mathbf{X}$ et $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ celles de $\mathbf{I}$. Les équations (4.13) sont l'expression matricielle de n systèmes de matrice associée commune $\mathbf{A}$ et de seconds membres $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$. Suivant 3.4.5, ces systèmes peuvent être résolus simultanément en réduisant la matrice augmentée n fois

$$\left[\begin{array}{c|c|c|c} \quad\mathbf{A}\quad & \mathbf{e}_1 & \dots & \mathbf{e}_n \end{array}\right] = \left[\begin{array}{c|c} \quad\mathbf{A}\quad & \quad\mathbf{I}\quad \end{array}\right]$$

à la forme échelonnée simplifiée. Compte tenu de 3.3.11, le résultat de cette réduction est une matrice de la forme

$$\left[\begin{array}{c|cccc} & b_{11} & b_{12} & \dots & b_{1n} \\ & b_{21} & b_{22} & \dots & b_{2n} \\ \mathbf{I} & & & & \\ & & \dots & \dots & \\ & b_{n1} & b_{n2} & \dots & b_{nn} \end{array}\right]. \tag{4.14}$$

Posons $\mathbf{B} = (b_{ij})$ et désignons par $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ les colonnes de $\mathbf{B}$. Les solutions des n systèmes réduits associés à la matrice augmentée (4.14) sont

$$\mathbf{x}_1 = \mathbf{b}_1,\ \mathbf{x}_2 = \mathbf{b}_2,\ \dots,\ \mathbf{x}_n = \mathbf{b}_n.$$

Par conséquent, $\mathbf{B}$ est une solution de l'équation (4.12). La proposition 4.2.3 nous permet alors de conclure que $\mathbf{A}$ est inversible et que $\mathbf{B} = \mathbf{A}^{-1}$.

Une méthode de calcul de la matrice inverse faisant appel aux déterminants sera présentée dans 5.3.6.

4.2.6 Proposition. Propriétés de l'opération d'inversion

Soit $\mathbf{A}$ *et* $\mathbf{B}$ *des matrices carrées du même ordre.*

(a) *Si* $\mathbf{A}$ *est inversible,* $\mathbf{A}^{-1}$ *et* ${}^t\mathbf{A}$ *le sont également et* $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$, $({}^t\mathbf{A})^{-1} = {}^t(\mathbf{A}^{-1})$.

(b) *Si* $\mathbf{A}$ *et* $\mathbf{B}$ *sont inversibles, le produit* $\mathbf{AB}$ *l'est également et* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

(c) *Si le produit* $\mathbf{AB}$ *est inversible,* $\mathbf{A}$ *et* $\mathbf{B}$ *le sont également.*

DÉMONSTRATION

Assertion (a). Par définition de l'inverse, $\mathbf{A}$ est la matrice inverse de $\mathbf{A}^{-1}$. D'autre part, en vertu de (d) de 4.1.9, ${}^t\mathbf{A}\,{}^t(\mathbf{A}^{-1}) = {}^t(\mathbf{A}^{-1}\mathbf{A}) = {}^t\mathbf{I} = \mathbf{I}$, donc, par la proposition 4.2.3, ${}^t(\mathbf{A}^{-1})$ est la matrice inverse de ${}^t\mathbf{A}$.

Assertion (b). La conclusion procède encore de la proposition 4.2.3, car $(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{I}$.

Assertion (c). On peut écrire $\mathbf{A}(\mathbf{B}(\mathbf{AB})^{-1}) = (\mathbf{AB})(\mathbf{AB})^{-1} = \mathbf{I}$, ainsi que $((\mathbf{AB})^{-1}\mathbf{A})\mathbf{B} = (\mathbf{AB})^{-1}(\mathbf{AB}) = \mathbf{I}$, d'où la conclusion s'ensuit, grâce à la proposition 4.2.3.

4.2.7 Corollaire. Inversion d'un produit de plusieurs facteurs et d'une puissance

Si $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ *sont des matrices inversibles du même ordre, le produit* $\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_k$ *est une matrice inversible et* $(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1}\mathbf{A}_{k-1}^{-1}\ldots\mathbf{A}_1^{-1}$. *En particulier, si* $\mathbf{A}$ *est une matrice inversible, la matrice* $\mathbf{A}^k$ *l'est également pour tout entier positif* k *et* $(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k$.

4.2.8 Remarques

(1) On désigne la matrice $(\mathbf{A}^{-1})^k$ plus simplement par $\mathbf{A}^{-k}$. La notion de puissance est ainsi étendue aux exposants entiers négatifs. A noter que lorsque $\mathbf{A}$ est inversible, les relations (4.2) sont vraies pour tous les entiers k et l; en outre, grâce à (a) de la proposition 4.2.6, la relation (f) de 4.1.9 s'étend à tous les entiers k.

(2) La somme de deux matrices inversibles n'est en général pas inversible. Par exemple, si $\mathbf{A}$ est inversible, $-\mathbf{A}$ l'est également, mais $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$ ne l'est pas.

4.2.9 Exemples

(1) La matrice-unité $\mathbf{I}$ est inversible et $\mathbf{I}^{-1} = \mathbf{I}$.

(2) On constate facilement que

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix}^{-1} = \begin{pmatrix} 3^{-1} & 0 & 0 \\ 0 & 2^{-1} & 0 \\ 0 & 0 & -4^{-1} \end{pmatrix}.$$

Cet exemple sera généralisé dans la section suivante.

(3) Nous allons calculer la matrice inverse de

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 3 & -1 \\ 2 & 1 & -5 \end{pmatrix}$$

par réduction de $(\mathbf{A} \vdots \mathbf{I})$ à la forme échelonnée simplifiée. Voici la suite des transformations:

$$\left(\begin{array}{cccccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 1 & 3 & -1 & 0 & 1 & 0 \\ 2 & 1 & -5 & 0 & 0 & 1 \end{array}\right)\begin{matrix} \\ \mathbf{L}_2 - \mathbf{L}_1 \\ \mathbf{L}_3 - 2\mathbf{L}_1 \end{matrix} \quad \left(\begin{array}{cccccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & -3 & -1 & -2 & 0 & 1 \end{array}\right)\begin{matrix} \mathbf{L}_1 - 2\mathbf{L}_2 \\ \\ \mathbf{L}_3 + 3\mathbf{L}_2 \end{matrix}$$

$$\left(\begin{array}{cccccc} 1 & 0 & -4 & 3 & -2 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 2 & -5 & 3 & 1 \end{array}\right)\begin{matrix} \mathbf{L}_1 + 2\mathbf{L}_3 \\ 2\mathbf{L}_2 - \mathbf{L}_3 \\ \\ \end{matrix} \quad \left(\begin{array}{cccccc} 1 & 0 & 0 & -7 & 4 & 2 \\ 0 & 2 & 0 & 3 & -1 & -1 \\ 0 & 0 & 2 & -5 & 3 & 1 \end{array}\right)\begin{matrix} \\ \frac{1}{2}\mathbf{L}_2 \\ \frac{1}{2}\mathbf{L}_3 \end{matrix}$$

La matrice inverse est donc

$$\mathbf{A}^{-1} = \tfrac{1}{2}\begin{pmatrix} -14 & 8 & 4 \\ 3 & -1 & -1 \\ -5 & 3 & 1 \end{pmatrix}.$$

(4) Nous nous proposons de calculer la matrice inverse de

$$\mathbf{A} = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix},$$

où a, b, c et d sont des nombres non tous nuls. Les colonnes de $\mathbf{A}$ sont orthogonales deux à deux et de même norme. Cela entraîne aussitôt que

$${}^t\mathbf{A}\mathbf{A} = (a^2 + b^2 + c^2 + d^2)\mathbf{I}_4.$$

Il en résulte que

$$\mathbf{A}^{-1} = \frac{1}{a^2 + b^2 + c^2 + d^2}\,{}^t\mathbf{A}.$$

(5) Soit $\mathbf{A}$ une matrice inversible d'ordre m, $\mathbf{B}$ une matrice inversible d'ordre n et $\mathbf{C}$ une matrice de type $m \times n$. Considérons l'équation matricielle

$$\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{pmatrix}\begin{pmatrix} \mathbf{X} & \mathbf{Z} \\ \mathbf{O} & \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix},$$

où $\mathbf{X}$, $\mathbf{Y}$ et $\mathbf{Z}$ sont des matrices inconnues. En multipliant par blocs et en égalant les sous-matrices du produit ainsi obtenu à celles du second membre, nous voyons que

$$\mathbf{AX} = \mathbf{I}_m, \quad \mathbf{BY} = \mathbf{I}_n, \quad \mathbf{AZ} + \mathbf{CY} = \mathbf{O}.$$

Les solutions de ces équations sont

$$\mathbf{X} = \mathbf{A}^{-1}, \quad \mathbf{Y} = \mathbf{B}^{-1}, \quad \mathbf{Z} = -\mathbf{A}^{-1}\mathbf{C}\mathbf{B}^{-1}.$$

Ainsi,

$$\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{C}\mathbf{B}^{-1} \\ \mathbf{O} & \mathbf{B}^{-1} \end{pmatrix}. \tag{4.15}$$

4.3 MATRICES CARRÉES PARTICULIÈRES

4.3.1 Matrices de type 1×1

Effectuées sur les matrices de type 1×1 (c'est-à-dire les matrices à un seul terme), les opérations définies dans 4.1.2, 4.1.3, 4.1.5 et 4.2.2 se confondent avec les opérations ordinaires sur les nombres. C'est pourquoi nous identifierons les matrices de ce type aux nombres.

Grâce à cette identification nous pourrons, par exemple, écrire le produit scalaire canonique dans $\mathbb{R}^n$ et la norme correspondante sous la forme

$$(\mathbf{x} \mid \mathbf{y}) = {}^t\mathbf{x}\mathbf{y}, \quad \| \mathbf{x} \| = \sqrt{{}^t\mathbf{x}\mathbf{x}}.$$

4.3.2 Matrices scalaires

On appelle *matrice scalaire* toute matrice de la forme

$$a\mathbf{I} = \begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & a & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & a \end{pmatrix}.$$

Nous avons déjà constaté que la multiplication d'une matrice scalaire $a\mathbf{I}$ par une matrice $\mathbf{A}$ est équivalente à la multiplication du nombre a par $\mathbf{A}$:

$$(a\mathbf{I})\mathbf{A} = a\mathbf{A}, \quad \mathbf{A}(a\mathbf{I}) = a\mathbf{A}.$$

4.3.3 Matrices diagonales

On appelle *matrice diagonale*, et on note $\operatorname{diag}(a_1, a_2, \ldots, a_n)$, toute matrice de la forme

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

Un calcul facile montre que

$$\operatorname{diag}(a_1, a_2, \ldots, a_n)\operatorname{diag}(b_1, b_2, \ldots, b_n) = \operatorname{diag}(a_1b_1, a_2b_2, \ldots, a_nb_n). \tag{4.16}$$

D'après la proposition 4.2.4, une matrice diagonale est inversible si et seulement si ses termes diagonaux sont tous non nuls. Si cette condition est remplie, il découle de (4.16) que

$$(\operatorname{diag}(a_1, a_2, \ldots, a_n))^{-1} = \operatorname{diag}(a_1^{-1}, a_2^{-1}, \ldots, a_n^{-1}).$$

La formule (4.16) s'étend, de manière évidente, au cas d'une famille finie quelconque de matrices diagonales. En particulier,

$$(\operatorname{diag}(a_1, a_2, \ldots, a_n))^k = \operatorname{diag}(a_1^k, a_2^k, \ldots, a_n^k), \tag{4.17}$$

où k est un entier quelconque ou un entier non négatif suivant que $\operatorname{diag}(a_1, a_2, \ldots, a_n)$ est inversible ou non.

Plus généralement, l'effet de la multiplication d'une matrice (non nécessairement carrée) par une matrice diagonale est le suivant:

$$\operatorname{diag}(a_1, a_2, \ldots, a_n) \begin{pmatrix} \mathbf{L}_1 \\ \hdashline \cdot \\ \cdot \\ \cdot \\ \hdashline \mathbf{L}_n \end{pmatrix} = \begin{pmatrix} a_1\mathbf{L}_1 \\ \hdashline \cdot \\ \cdot \\ \cdot \\ \hdashline a_n\mathbf{L}_n \end{pmatrix},$$

$$\left(\begin{array}{c:c:c} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{array}\right) \operatorname{diag}(a_1, a_2, \ldots, a_n) = \left(\begin{array}{c:c:c} a_1\mathbf{b}_1 & \cdots & a_n\mathbf{b}_n \end{array}\right).$$

4.3.4 Matrices triangulaires

On appelle *matrice triangulaire supérieure* toute matrice de la forme

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

On appelle *matrice triangulaire inférieure* la transposée d'une matrice triangulaire supérieure.

Pour qu'une matrice triangulaire supérieure (inférieure) $\mathbf{A}$ soit inversible il faut et il suffit que ses termes diagonaux $a_{11}, a_{22}, \ldots, a_{nn}$ soient tous non nuls. Cette condition étant remplie, $\mathbf{A}^{-1}$ est encore une matrice triangulaire supérieure (inférieure); en outre, les termes diagonaux de $\mathbf{A}^{-1}$ sont $a_{11}^{-1}, a_{22}^{-1}, \ldots, a_{nn}^{-1}$. Pour démontrer ces assertions, il suffit d'observer l'effet des opérations élémentaires sur la matrice $(\mathbf{A} \,\vdots\, \mathbf{I})$, lorsqu'on la transforme en une matrice échelonnée simplifiée. On peut également procéder par récurrence sur n en s'appuyant sur la formule (4.15).

4.3.5 Matrices symétriques

On appelle *matrice symétrique* toute matrice égale à sa transposée. Entre les termes a_{ij} d'une matrice symétrique il y a donc la relation

$$a_{ij} = a_{ji}. \tag{4.18}$$

Toute matrice diagonale est évidemment symétrique. Voici un exemple de matrice symétrique non diagonale:

$$\begin{bmatrix} -1 & 2 & 5 \\ 2 & 0 & -1 \\ 5 & -1 & 3 \end{bmatrix}.$$

L'inverse d'une matrice symétrique inversible est encore une matrice symétrique, car, d'après (a) de la proposition 4.2.6,

$${}^t(\mathbf{A}^{-1}) = ({}^t\mathbf{A})^{-1} = \mathbf{A}^{-1}.$$

Si **A** et **B** sont des matrices symétriques du même ordre, le produit **AB** n'est symétrique que si **A** et **B** commutent. En effet, d'après la propriété (d) de 4.1.9,

$${}^t(\mathbf{AB}) = {}^t\mathbf{B}\,{}^t\mathbf{A} = \mathbf{BA}.$$

4.3.6 Remarque

L'addition et la multiplication par un scalaire de matrices de l'une quelconque des classes présentées dans cette section est une matrice de la même classe. En d'autres termes, les matrices d'ordre n de chaque classe forment un sous-espace vectoriel de l'espace vectoriel des matrices carrées d'ordre n.

4.4 RETOUR AUX OPÉRATIONS ÉLÉMENTAIRES

4.4.1 Matrices élémentaires

On appelle *matrice élémentaire* toute matrice pouvant se déduire de la matrice-unité par l'une des opérations élémentaires définies dans 3.3.4. Toute matrice carrée d'ordre 1 est donc une matrice élémentaire et toute matrice élémentaire d'ordre supérieur à 1 est de l'un des trois types suivants:

Type 1: $\mathbf{I}_{il}$ $(i \neq l)$ se déduit de **I** par échange de la i-ième et la l-ième ligne; par exemple, si $i < l$,

$$\mathbf{I}_{il} = \begin{array}{l} \begin{pmatrix} 1 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 1 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 1 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 1 \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow i\text{-ième} \\ \\ \leftarrow l\text{-ième} \\ \\ \\ \end{array} \\ \qquad\quad \uparrow \qquad \uparrow \\ \qquad i\text{-ième}\ l\text{-ième} \end{array}$$

Type 2: $\mathbf{I}_{il}^{a}$ $(i \neq l)$ se déduit de **I** par addition à la i-ième ligne du produit de la l-ième ligne par le nombre a; par exemple, si $i < l$,

$$\mathbf{I}_{il}^{a} = \begin{array}{l} \begin{pmatrix} 1 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 1 & \cdot\,\cdot & a & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 1 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 1 \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow i\text{-ième} \\ \\ \leftarrow l\text{-ième} \\ \\ \\ \end{array} \\ \qquad\quad \uparrow \qquad \uparrow \\ \qquad i\text{-ième}\ l\text{-ième} \end{array}$$

Type 3: $\mathbf{I}_{i}^{a}$ se déduit de **I** par multiplication de la i-ième ligne par le nombre a; autrement dit,

$$\mathbf{I}_{i}^{a} = \begin{pmatrix} 1 & 0 & \cdots & & & & 0 \\ 0 & 1 & \cdots & & & & 0 \\ & & \ddots & & & & \\ & & & 1 & & & \\ \cdots & & & & a & & \cdots \\ & & & & & 1 & \\ & & & & & & \ddots \\ 0 & 0 & \cdots & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \\ \leftarrow i\text{-ième} \\ \\ \\ \\ \end{array}$$

4.4.2 Effet de la multiplication par une matrice élémentaire

L'intérêt des matrices élémentaires vient du fait qu'elles permettent d'exprimer les opérations de réduction d'une matrice par des multiplications matricielles. Supposons que $\mathbf{I}_{il}$, $\mathbf{I}_{il}^{a}$ et $\mathbf{I}_{i}^{a}$ soient d'ordre n. Alors, pour toute matrice **A** à n lignes:

(a) $\mathbf{I}_{il}\mathbf{A}$ se déduit de **A** par échange de la i-ième et la l-ième ligne.

(b) $\mathbf{I}_{il}^{a}\mathbf{A}$ se déduit de **A** par addition à la i-ième ligne du produit de la l-ième ligne par a.

(c) $\mathbf{I}_{i}^{a}\mathbf{A}$ se déduit de **A** par multiplication de la i-ième ligne par a.

4.4.3 Transposée et inverse d'une matrice élémentaire

La transposée et l'inverse d'une matrice élémentaire sont encore des matrices élémentaires. Plus précisément:

(a) ${}^t\mathbf{I}_{il} = \mathbf{I}_{il}$, $\mathbf{I}_{il}^{-1} = \mathbf{I}_{il}$.
(b) ${}^t(\mathbf{I}_{il}^a) = \mathbf{I}_{li}^a$, $(\mathbf{I}_{il}^a)^{-1} = \mathbf{I}_{il}^{-a}$.
(c) ${}^t(\mathbf{I}_i^a) = \mathbf{I}_i^a$, $(\mathbf{I}_i^a)^{-1} = \mathbf{I}_i^{1/a}$ $(a \neq 0)$.

4.4.4 Matrices inversibles comme produits de matrices élémentaires

Selon 3.3.5, 3.3.7 et 4.4.2, toute matrice peut être transformée en une matrice échelonnée ou en une matrice échelonnée simplifiée au moyen d'une suite finie de multiplications par une matrice élémentaire. En particulier, pour toute matrice inversible $\mathbf{A}$, il existe, vu 3.3.11, des matrices élémentaires $\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_k$ telles que

$$\mathbf{M}_1\mathbf{M}_2 \ldots \mathbf{M}_k\mathbf{A} = \mathbf{I}.$$

En multipliant les deux membres par $(\mathbf{M}_1\mathbf{M}_2 \ldots \mathbf{M}_k)^{-1}$, nous en concluons, grâce au corollaire 4.2.7, que

$$\mathbf{A} = (\mathbf{M}_1\mathbf{M}_2 \ldots \mathbf{M}_k)^{-1} = \mathbf{M}_k^{-1}\mathbf{M}_{k-1}^{-1} \ldots \mathbf{M}_1^{-1},$$

ce qui atteste que toute matrice inversible est un produit de matrices élémentaires.

4.5 FONCTIONS MATRICIELLES

4.5.1 Polynômes matriciels

Si $\mathbf{A}$ est une matrice carrée et $f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_k x^k$ un polynôme, nous poserons

$$f(\mathbf{A}) = \alpha_0\mathbf{I} + \alpha_1\mathbf{A} + \ldots + \alpha_k\mathbf{A}^k. \tag{4.19}$$

La fonction qui associe à toute matrice carrée $\mathbf{A}$ la matrice $f(\mathbf{A})$ est appelée *polynôme matriciel.*

4.5.2 Fonctions matricielles

Nous pouvons définir des fonctions matricielles plus générales que les polynômes à l'aide de séries matricielles. Soit $\mathbf{A}$ une matrice carrée d'ordre n. Nous dirons qu'une série

$$\sum_{k=0}^{\infty} \alpha_k \mathbf{A}^k \tag{4.20}$$

converge (*absolument*) si chacun des n^2 termes de la matrice

$$\sum_{k=0}^{l} \alpha_k \mathbf{A}^k$$

converge (absolument) quand l tend vers l'infini. Cette condition étant remplie, nous appellerons *somme* de la série (4.20) la matrice formée des termes-limites. Cette somme est encore désignée par l'expression (4.20).

A noter que si la matrice $\mathbf{A}$ est nilpotente, la série (4.20) se réduit à une somme finie, car $\mathbf{A}^k = \mathbf{O}$ pour tout entier k assez grand.

Supposons qu'une fonction réelle f soit définie par une série entière dans un voisinage de 0:

$$f(x) = \sum_{k=0}^{\infty} \alpha_k x^k, \quad |x| < r. \tag{4.21}$$

La série (4.20) converge alors absolument pour toute matrice carrée $\mathbf{A}$ d'ordre n telle que $\|\mathbf{A}\| < r$ (cf. exercices 4.7.19 et 4.7.18), où

$$\|\mathbf{A}\| = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}. \tag{4.22}$$

Pour une telle matrice $\mathbf{A}$ nous poserons

$$f(\mathbf{A}) = \sum_{k=0}^{\infty} \alpha_k \mathbf{A}^k. \tag{4.23}$$

Dans le cas où $\mathbf{A}$ est diagonale, cette définition de $f(\mathbf{A})$ se réduit, grâce à (4.17), à

$$f(\mathrm{diag}(a_1, a_2, \ldots, a_n)) = \mathrm{diag}(f(a_1), f(a_2), \ldots, f(a_n)). \tag{4.24}$$

4.5.3 Dérivation de fonctions matricielles

On dit qu'une matrice dépendant d'un paramètre réel t est *dérivable* si tous ses termes sont dérivables. Soit $\mathbf{A}(t) = (a_{ij}(t))$ une matrice dérivable. On appelle *dérivée* de $\mathbf{A}(t)$ la matrice $\dot{\mathbf{A}}(t) = (\dot{a}_{ij}(t))$, où $\dot{a}_{ij}(t)$ désigne la dérivée de $a_{ij}(t)$.

Les seules matrices dérivables qui nous intéressent se déduisent de (4.23) par substitution de $t\mathbf{A}$ à $\mathbf{A}$:

$$f(t\mathbf{A}) = \sum_{k=0}^{\infty} \alpha_k t^k \mathbf{A}^k.$$

D'après ce qui a été dit dans 4.5.2, cette série converge absolument pour tout t tel que

$$|t| < \frac{r}{\|\mathbf{A}\|},$$

où r est la borne apparaissant dans (4.21) (par convention $r/0 = \infty$). En outre, $f(t\mathbf{A})$ est dérivable et

$$(f(t\mathbf{A}))^{\boldsymbol{\cdot}} = \mathbf{A}\dot{f}(t\mathbf{A}), \tag{4.25}$$

où le point dans le premier membre indique la dérivation par rapport à t de la matrice entre parenthèses et $\dot{f}$ dans le second membre désigne la dérivée de f. Dans le cas où f est un polynôme, cette relation résulte des calculs suivants:

$$\begin{aligned}(f(t\mathbf{A}))^{\cdot} &= (\alpha_0\mathbf{I} + \alpha_1 t\mathbf{A} + \alpha_2 t^2\mathbf{A}^2 + \dots + \alpha_k t^k\mathbf{A}^k)^{\cdot} \\ &= \alpha_1\mathbf{A} + 2\alpha_2 t\mathbf{A}^2 + \dots + k\alpha_k t^{k-1}\mathbf{A}^k \\ &= \mathbf{A}(\alpha_1\mathbf{I} + 2\alpha_2 t\mathbf{A} + \dots + k\alpha_k t^{k-1}\mathbf{A}^{k-1}) = \mathbf{A}\dot{f}(t\mathbf{A}).\end{aligned}$$

Le cas général s'ensuit par passage à la limite.

4.5.4 Exponentielle d'une matrice

L'*exponentielle* d'une matrice carrée $\mathbf{A}$ s'obtient à partir de (4.21) en posant $\alpha_k = 1/k!$ (dans ce cas, $r = \infty$):

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{A}^k. \tag{4.26}$$

Il est évident que $\exp(\mathbf{O}) = \mathbf{I}$ et $\exp(a\mathbf{I}) = e^a\mathbf{I}$. En outre, si $\mathbf{A}$ et $\mathbf{B}$ commutent

$$\exp(\mathbf{A} + \mathbf{B}) = \exp(\mathbf{A})\exp(\mathbf{B}). \tag{4.27}$$

En effet, compte tenu de (4.3),

$$\begin{aligned}\exp(\mathbf{A})\exp(\mathbf{B}) &= \left(\sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{A}^k\right)\left(\sum_{l=0}^{\infty}\frac{1}{l!}\mathbf{B}^l\right) = \sum_{l=0}^{\infty}\left(\sum_{k=0}^{l}\frac{1}{k!(l-k)!}\mathbf{A}^k\mathbf{B}^{l-k}\right) \\ &= \sum_{l=0}^{\infty}\frac{1}{l!}(\mathbf{A} + \mathbf{B})^l = \exp(\mathbf{A} + \mathbf{B}).\end{aligned}$$

En particulier,

$$\exp(\mathbf{A})\exp(-\mathbf{A}) = \exp(\mathbf{A} - \mathbf{A}) = \exp(\mathbf{O}) = \mathbf{I},$$

d'où il résulte que $\exp(\mathbf{A})$ est inversible et

$$(\exp(\mathbf{A}))^{-1} = \exp(-\mathbf{A}). \tag{4.28}$$

Appliquée à la fonction exponentielle, la relation (4.25) devient

$$(\exp(t\mathbf{A}))^{\cdot} = \mathbf{A}\exp(t\mathbf{A}). \tag{4.29}$$

4.6 MATRICES DE TRANSITION

4.6.1 Chaînes de Markov finies

Imaginons qu'une particule occupe une parmi n positions possibles, désignées par 1, 2, ..., n et appelées états. Supposons qu'aux instants 0, 1, 2, ... la particule change d'état de manière aléatoire. Ce changement est gouverné par des

probabilités appelées *probabilités de transition*. La suite des états successifs occupés par la particule au cours du temps constitue ce que l'on appelle un processus stochastique. Si la probabilité de transition d'un état i à un état j ne dépend que de i et de j, et non pas d'autres états ou de l'instant auquel le changement d'état a lieu, on dit que le processus stochastique est une *chaîne de Markov homogène* ou simplement une *chaîne de Markov*.

4.6.2 Matrices de transition

Etant donné une chaîne de Markov, désignons par a_{ij} la probabilité de transition de l'état i à l'état j en une unité de temps. On appelle la matrice $\mathbf{A} = (a_{ij})$ *matrice de transition* de la chaîne. Cette matrice satisfait aux conditions suivantes:

$$\begin{aligned}&0 \leqslant a_{ij} \leqslant 1, \quad i, j = 1, 2, \dots, n, \\ &a_{i1} + a_{i2} + \dots + a_{in} = 1, \quad i = 1, 2, \dots, n.\end{aligned}$$

La première condition exprime simplement le fait qu'une probabilité se mesure par un nombre compris entre 0 et 1; la deuxième que la chaîne passe, avec certitude, de i à un des états 1, 2, ..., n.

4.6.3 Loi initiale

Supposons que l'état initial de la chaîne ne soit pas fixe, mais gouverné, lui aussi, par des probabilités: $p_j^{(0)}$ est la probabilité qu'à l'instant 0 la chaîne occupe l'état j. Ces probabilités satisfont aux conditions suivantes:

$$\begin{aligned}&0 \leqslant p_j^{(0)} \leqslant 1, \quad j = 1, 2, \dots, n, \\ &p_1^{(0)} + p_2^{(0)} + \dots + p_n^{(0)} = 1.\end{aligned}$$

On appelle la matrice-ligne

$$\mathbf{L}^{(0)} = \begin{bmatrix} p_1^{(0)} & p_2^{(0)} & \dots & p_n^{(0)} \end{bmatrix}$$

loi initiale de la chaîne.

4.6.4 Loi à l'instant k

La loi initiale et la matrice de transition d'une chaîne de Markov déterminent la probabilité de tout événement lié à cette chaîne. Par exemple, désignons par $p_j^{(k)}$ la probabilité qu'à l'instant k la chaîne occupe l'état j. On appelle la matrice-ligne

$$\mathbf{L}^{(k)} = \begin{bmatrix} p_1^{(k)} & p_2^{(k)} & \dots & p_n^{(k)} \end{bmatrix}$$

loi de la chaîne à l'instant k. Il est intuitivement clair que

$$p_j^{(1)} = p_1^{(0)}a_{1j} + p_2^{(0)}a_{2j} + \dots + p_n^{(0)}a_{nj}$$

et, plus généralement, pour tout instant positif k, que

$$p_j^{(k)} = p_1^{(k-1)}a_{1j} + p_2^{(k-1)}a_{2j} + \dots + p_n^{(k-1)}a_{nj}.$$

En termes matriciels cette relation s'écrit

$$\mathbf{L}^{(k)} = \mathbf{L}^{(k-1)}\mathbf{A}.$$

Par récurrence, il s'ensuit que

$$\mathbf{L}^{(k)} = \mathbf{L}^{(0)}\mathbf{A}^k. \tag{4.30}$$

Nous voyons ainsi que la loi initiale et la matrice de transition permettent de calculer la loi de la chaîne à n'importe quel instant.

4.6.5 Exemple

Un magasin vend deux marques m_1 et m_2 d'un certain produit et un client habituel de la maison achète m_1 avec probabilité 0,7 si son dernier achat était m_1 et avec probabilité 0,4 si son dernier achat était m_2. Nous allons calculer les probabilités d'achat de m_1 et de m_2 lors du sixième achat (l'achat initial est fait à l'instant 0). Nous avons à faire à une chaîne de Markov à deux états, l'état 1 désignant la marque m_1 et l'état 2 la marque m_2. La matrice de transition de cette chaîne est

$$\mathbf{A} = \begin{bmatrix} 0{,}7 & 0{,}3 \\ 0{,}4 & 0{,}6 \end{bmatrix}.$$

Supposons, par exemple, que la loi initiale soit

$$\mathbf{L}^{(0)} = \begin{bmatrix} \tfrac{1}{3} & \tfrac{2}{3} \end{bmatrix}.$$

La loi à l'instant 5 est alors, d'après (4.30),

$$\mathbf{L}^{(5)} = \mathbf{L}^{(0)}\mathbf{A}^5 \cong \begin{bmatrix} \tfrac{1}{3} & \tfrac{2}{3} \end{bmatrix} \begin{bmatrix} 0{,}57247 & 0{,}42753 \\ 0{,}57004 & 0{,}42996 \end{bmatrix} \cong \begin{bmatrix} 0{,}57084 & 0{,}42914 \end{bmatrix}.$$

Lors du sixième achat, les probabilités d'achat de m_1 et de m_2 sont donc respectivement 0,57084 et 0,42914.

On peut se demander si l'influence du choix initial tend à disparaître après un grand nombre d'achats. Nous répondrons à cette question plus loin, lorsque nous disposerons des moyens nécessaires pour étudier le comportement asymptotique des puissances de certaines matrices (cf. 8.4.5).

4.7 EXERCICES

4.7.1 Soit

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 3 \\ 3 & 2 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & 1 & 4 \\ 1 & 2 & -1 \\ -1 & 1 & 2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Calculer, lorsque la multiplication est possible, les produits **AA**, **BB**, **AB**, **AC**, **BC**, ${}^t\mathbf{C}\mathbf{A}$, ${}^t\mathbf{C}\mathbf{B}$, ${}^t\mathbf{C}\mathbf{C}$, $\mathbf{C}\,{}^t\mathbf{C}$.

4.7.2 Soit $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ les colonnes d'une matrice **A** de type $m \times n$ et $\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_n$ les lignes d'une matrice **B** de type $n \times p$. Montrer que

$$\mathbf{AB} = \mathbf{a}_1\mathbf{B}_1 + \mathbf{a}_2\mathbf{B}_2 + \dots + \mathbf{a}_n\mathbf{B}_n.$$

4.7.3 Par réduction simultanée, trouver l'ensemble des solutions de l'équation matricielle $\mathbf{AX} = \mathbf{B}$, où

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{et} \quad \mathbf{B} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \\ 5 & 2 \end{bmatrix}.$$

4.7.4 Ecrire la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

sous la forme $\mathbf{I} + \mathbf{B}$ et calculer $\mathbf{A}^k$, pour tout nombre entier positif k, à l'aide de la formule du binôme.

4.7.5 Montrer que les seules matrices qui commutent avec la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & 0 & \\ & & 0 & 1 & & \\ & & & \ddots & \ddots & \\ & 0 & & & \ddots & 1 \\ & & & & & 0 \end{bmatrix} \quad \text{sont celles de la forme} \quad \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ & a_1 & a_2 & \cdots & a_{n-1} \\ & & a_1 & \cdots & a_{n-2} \\ & 0 & & \ddots & \vdots \\ & & & & a_1 \end{bmatrix}.$$

En déduire que les matrices qui commutent avec **A** et ${}^t\mathbf{A}$ sont de la forme $a\mathbf{I}$. Relever, au passage, que si deux matrices commutent, aucune des deux, en général, ne commute avec la transposée de l'autre.

4.7.6 Par récurrence sur k, démontrer que la formule du binôme (4.3) s'applique à tout couple $(\mathbf{A}, \mathbf{B})$ de matrices carrées qui commutent.

4.7.7 Soit **A** une matrice de type $m \times n$ et de rang n.

(a) Montrer que le rang de ${}^t\mathbf{AA}$ est n. (Multiplier les deux membres de l'équation ${}^t\mathbf{AAx} = \mathbf{0}$ pat ${}^t\mathbf{x}$, en déduire que $\mathbf{Ax} = \mathbf{0}$ et appliquer le corollaire 3.5.4.)

(b) En déduire que la seule solution de toute équation de la forme $\mathbf{AX} = \mathbf{O}$ est $\mathbf{X} = \mathbf{O}$.

4.7.8 *Méthode des moindres carrés.* Soit $P_1(a_1, b_1)$, $P_2(a_2, b_2)$, ..., $P_n(a_n, b_n)$ n points de $\mathbb{R}^2$ tels que $a_i \neq a_j$ si $i \neq j$. Soit k un entier positif inférieur à n. Pour tout polynôme p de degré inférieur ou égal à k, on pose

$$\delta^2(p) = (b_1 - p(a_1))^2 + (b_2 - p(a_2))^2 + ... + (b_n - p(a_n))^2.$$

Le but de cet exercice est de trouver le polynôme $p(t) \equiv \alpha_0 + \alpha_1 t + ... + \alpha_k t^k$ qui minimise $\delta(p)$. On désigne par **p** le vecteur-colonne formé des coefficients inconnus $\alpha_0, \alpha_1, ..., \alpha_k$ et on pose

$$\mathbf{A} = \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^k \\ 1 & a_2 & a_2^2 & \cdots & a_2^k \\ & & \cdots\cdots & & \\ 1 & a_n & a_n^2 & \cdots & a_n^k \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Les assertions suivantes sont à démontrer:

(a) $\delta(p) = \| \mathbf{b} - \mathbf{Ap} \|$.

(b) $\delta(p)$ est minimal lorsque **Ap** est la projection orthogonale de **b** sur le sous-espace vectoriel formé des vecteurs-colonnes **Ax**, **x** parcourant $\mathbb{R}^{k+1}$.

(c) **Ap** est cette projection lorsque **p** est l'unique solution de l'équation

${}^t\mathbf{AAp} = {}^t\mathbf{Ab}$.

(Poser que $\mathbf{b} - \mathbf{Ap}$ est orthogonal à **Ax** pour tout **x** de $\mathbb{R}^{k+1}$; utiliser les exercices 3.6.6 et 4.7.7 pour prouver l'unicité de la solution.)

Application numérique: étant donné $P_1(0, 1)$, $P_2(1, 2)$, $P_3(2, 4)$, $P_4(3, 8)$, trouver le polynôme p de degré inférieur ou égal à 2 qui minimise $\delta(p)$.

4.7.9 Montrer que si **A** est une matrice inversible, $\alpha\mathbf{A}$ l'est également pour tout nombre α non nul et $(\alpha\mathbf{A})^{-1} = \frac{1}{\alpha}\mathbf{A}^{-1}$.

4.7.10 Montrer que si deux matrices inversibles **A** et **B** commutent, alors $\mathbf{A}^{-1}$ et $\mathbf{B}^{-1}$ commutent également.

4.7.11 Montrer que si une matrice inversible **A** commute avec une matrice **B**, alors $\mathbf{A}^{-1}$ commute avec **B**.

4.7.12 Soit **A** une matrice carrée. Soit **B** et **C** des matrices inversibles du même ordre que **A** telles que ${}^t(\mathbf{BA})\mathbf{B}^{-1}(\mathbf{CB}^{-1})^{-1} = \mathbf{I}$. Montrer que **A** est inversible et exprimer $\mathbf{A}^{-1}$ en fonction de **B** et **C**. (Appliquer la proposition 4.2.3.)

4.7.13 Soit **A** une matrice carrée telle que $\mathbf{A}^3 - \mathbf{A}^2 + 2\mathbf{A} - 3\mathbf{I} = \mathbf{O}$. Montrer que **A** est inversible et exprimer $\mathbf{A}^{-1}$ sous la forme d'une combinaison linéaire de puissances entières non négatives de **A**. (Appliquer la proposition 4.2.3.)

4.7.14 Soit **A** une matrice inversible d'ordre m, **B** une matrice inversible d'ordre n et **C** une matrice de type $m \times n$. Montrer que

$$\begin{bmatrix} \mathbf{C} & \mathbf{A} \\ \mathbf{B} & \mathbf{O} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{O} & \mathbf{B}^{-1} \\ \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{CB}^{-1} \end{bmatrix}.$$

A l'aide de cette formule, calculer la matrice inverse de

$$\begin{bmatrix} -1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

4.7.15 Calculer la matrice inverse de

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ 2 & 0 & 2 & 3 \\ 2 & -5 & 1 & 0 \\ 1 & -2 & -2 & 2 \end{bmatrix}.$$

4.7.16 On appelle *matrice antisymétrique* toute matrice égale à sa transposée multipliée par -1.

(a) Montrer que pour toute matrice carrée **A**, $\mathbf{A} + {}^t\mathbf{A}$ est une matrice symétrique et $\mathbf{A} - {}^t\mathbf{A}$ une matrice antisymétrique.

(b) Montrer que l'ensemble des matrices symétriques d'ordre n et celui des matrices antisymétriques d'ordre n sont des sous-espaces complémentaires dans l'espace vectoriel des matrices carrées d'ordre n.

(c) Trouver la dimension de ces deux sous-espaces et vérifier la relation (1.12).

4.7.17 Soit **D** une matrice diagonale d'ordre n dont les termes diagonaux sont distincts. Démontrer que toute matrice diagonale d'ordre n s'écrit de manière unique sous la forme d'une combinaison linéaire des matrices $\mathbf{I}, \mathbf{D}, \mathbf{D}^2, ..., \mathbf{D}^{n-1}$. (Utiliser l'exercice 3.6.6.)

4.7.18 Dans l'espace vectoriel des matrices carrées d'ordre n, on considère l'opération $(\cdot \mid \cdot)$ définie par la formule

$(\mathbf{A} \mid \mathbf{B}) = \mathrm{tr}({}^t\mathbf{AB})$.

(a) Montrer que cette opération est un produit scalaire.

(b) Montrer que la norme associée à ce produit scalaire est celle qui est définie dans (4.22).

(c) Montrer que

$$\| \mathbf{AB} \| \leqslant \| \mathbf{A} \| \| \mathbf{B} \|.$$

(d) Montrer que si $\mathbf{x}$ est un vecteur-colonne de $\mathbb{R}^n$,

$$\| \mathbf{Ax} \| \leqslant \| \mathbf{A} \| \| \mathbf{x} \|.$$

4.7.19 Montrer que la série (4.20) converge absolument si $\| \mathbf{A} \| < r$, où r est la borne qui apparaît sous (4.21).

4.7.20 Soit $\mathbf{A}$ une matrice carrée telle que $\| \mathbf{A} \| < 1$. Montrer que

$$(\mathbf{I} - \mathbf{A})^{-1} = \sum_{k=0}^{\infty} \mathbf{A}^k.$$

(Utiliser l'exercice précédent.)

4.7.21 Montrer que si

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{alors} \quad \exp(t\mathbf{A}) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

(Utiliser les développements en série de Taylor des fonctions cos et sin.)

CHAPITRE 5

Déterminants

5.1 DÉFINITION ET PROPRIÉTÉS DES DÉTERMINANTS

5.1.1 Introduction

La notion de déterminant est apparue en relation avec la résolution de certains systèmes linéaires. Elle est également liée à l'idée de volume, comme le montrent les conclusions de 2.7.14 et 2.7.15. Dans cette section, nous nous proposons d'exposer et de commenter les propriétés principales dérivant de la définition de cette nouvelle notion, ainsi que d'étendre le concept de volume à un parallélépipède n-dimensionnel. Dans la section 5.3, nous étudierons les développements et appliquerons les résultats obtenus à la résolution de certains systèmes linéaires, ainsi qu'au calcul de l'inverse d'une matrice.

5.1.2 Notation

Le *déterminant d'une matrice* carrée $\mathbf{A} = (a_{ij})$ est un nombre associé à $\mathbf{A}$ que nous nous apprêtons à définir et noterons

$$|\mathbf{A}| \quad \text{ou} \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

ou aussi

$$\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n),$$

lorsque la dépendance de $\mathbf{A}$ est exprimée au moyen de ses colonnes $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. En accord avec ce dernier symbole, le déterminant est également appelé *déterminant des vecteurs-colonnes* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$.

Par la suite, il sera souvent question de colonnes ou de lignes d'un déterminant. Il est entendu que ces colonnes ou ces lignes sont celles de la matrice à laquelle le déterminant est associé.

Si $\mathbf{A}$ est une matrice carrée d'ordre n supérieur à 1, nous désignerons par $\mathbf{A}_{ij}$ la matrice carrée d'ordre $n-1$ obtenue en supprimant la i-ième ligne et la j-ième colonne de $\mathbf{A}$.

5.1.3 Déterminant d'ordre n

On appelle *déterminant d'ordre n* la fonction qui associe à toute matrice carrée d'ordre n son déterminant. Nous définissons cette fonction par récurrence sur n:

- le déterminant d'une matrice carrée $\mathbf{A} = (a)$ d'ordre 1 est a;
- supposons que le déterminant d'ordre $n-1$ ait été défini pour un entier n supérieur à 1; le déterminant d'une matrice carrée $\mathbf{A} = (a_{ij})$ d'ordre n est alors le nombre

$$\begin{aligned} |\mathbf{A}| &= a_{11}|\mathbf{A}_{11}| - a_{21}|\mathbf{A}_{21}| + \dots + (-1)^{n+1}a_{n1}|\mathbf{A}_{n1}| \\ &= \sum_{i=1}^{n}(-1)^{i+1}a_{i1}|\mathbf{A}_{i1}|. \end{aligned} \tag{5.1}$$

On notera que pour $n = 2$ et $n = 3$, notre définition coïncide avec celle introduite dans 2.7.2:

$$n = 2: \ |\mathbf{A}| = a_{11}|(a_{22})| - a_{21}|(a_{12})| = a_{11}a_{22} - a_{21}a_{12};$$

$$\begin{aligned} n = 3: \ |\mathbf{A}| &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ &\quad - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}. \end{aligned}$$

5.1.4 Exemples

(1) Exemple de calcul d'un déterminant:

$$\begin{vmatrix} 0 & 1 & 1 & 2 \\ 3 & 2 & 0 & 1 \\ 0 & 1 & 2 & 4 \\ 2 & 3 & 2 & 0 \end{vmatrix} = -3\begin{vmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 3 & 2 & 0 \end{vmatrix} - 2\begin{vmatrix} 1 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 4 \end{vmatrix}$$

$$= -3(1\begin{vmatrix} 2 & 4 \\ 2 & 0 \end{vmatrix} - 1\begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} + 3\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix}) - 2(1\begin{vmatrix} 0 & 1 \\ 2 & 4 \end{vmatrix} - 2\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} + 1\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix})$$

$$= -3(-8 + 4 + 0) - 2(-2 - 0 + 1) = 14.$$

(2) *Déterminant d'une matrice diagonale*. Par application répétée de (5.1), nous obtenons:

$$\begin{vmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ & \dots\dots & & \\ 0 & 0 & \dots & a_n \end{vmatrix} = a_1\begin{vmatrix} a_2 & 0 & \dots & 0 \\ 0 & a_3 & \dots & 0 \\ & \dots\dots & & \\ 0 & 0 & \dots & a_n \end{vmatrix} = \dots = a_1a_2 \dots a_n. \tag{5.2}$$

En particulier,

$$|a\mathbf{I}_n| = a^n. \tag{5.3}$$

(3) *Déterminant d'une matrice triangulaire*. Par application répétée de (5.1), nous obtenons:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ & \dots\dots & & \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ & \dots\dots & & \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = \dots = a_{11}a_{22} \dots a_{nn}. \tag{5.4}$$

Nous verrons dans la proposition 5.1.8 que le déterminant d'une matrice carrée $\mathbf{A}$ est égal au déterminant de sa transposée ${}^t\mathbf{A}$. Nous en déduisons que le déterminant d'une matrice triangulaire inférieure est aussi égal au produit des termes diagonaux.

5.1.5 Théorème. Propriétés fondamentales du déterminant d'ordre n

Le déterminant d'ordre n jouit des propriétés suivantes:

(a) $\det(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots, \mathbf{a}_n) = -\det(\mathbf{a}_1, \dots, \mathbf{a}_k, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n)$ $(n > 1, j < k)$.
(b) $\det(\mathbf{a}_1, \dots, \alpha\mathbf{a}_j, \dots, \mathbf{a}_n) = \alpha\det(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n)$.
(c) $\det(\mathbf{a}_1, \dots, \mathbf{a}_j + \mathbf{a}'_j, \dots, \mathbf{a}_n) = \det(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n) + \det(\mathbf{a}_1, \dots, \mathbf{a}'_j, \dots, \mathbf{a}_n)$.
(d) $\det(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$, *où* $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ *est la base canonique de* $\mathbb{R}^n$.

En d'autres termes, le déterminant est multiplié par -1 si deux de ses colonnes sont échangées, il est multiplié par α si une de ses colonnes est multipliée par α, il est additif en chacune de ses colonnes et attribue la valeur 1 à la matrice-unité.

L'union des propriétés (a), (b) et (c) s'énonce en disant que le déterminant d'ordre n est une *forme n-linéaire alternée* dans $\mathbb{R}^n$ (cf. 6.4.3).

Le déterminant d'ordre n n'est pas additif (sauf si $n = 1$); autrement dit, $|\mathbf{A} + \mathbf{B}|$ n'est en général pas égal à $|\mathbf{A}| + |\mathbf{B}|$. Par exemple,

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \neq \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix}.$$

Nous verrons dans 5.2.4 que les propriétés (a)–(d) caractérisent le déterminant d'ordre n. Ces propriétés en entraînent d'autres, que nous énoncerons dans les deux propositions suivantes et démontrerons, avec le théorème 5.1.5, dans la section 5.2.

On remarquera d'ores et déjà que le déterminant est nul si deux de ses colonnes sont égales ou si l'une d'entre elles est nulle. En effet, un échange de deux colonnes égales laisse le déterminant inchangé et, d'après (a), change en même temps ce déterminant en son opposé, ce qui démontre la première assertion; quant à la seconde, il suffit de poser $\alpha = 0$ dans (b).

On notera en outre que

$$|\alpha\mathbf{A}| = \alpha^n |\mathbf{A}| \tag{5.5}$$

pour toute matrice carrée $\mathbf{A}$ d'ordre n. En effet, cela résulte de (b) appliquée à chaque colonne de $\mathbf{A}$.

5.1.6 Proposition. Propriétés du déterminant d'ordre n

Le déterminant d'ordre n jouit des propriétés suivantes:

(a) $\det(\mathbf{a}_1, \ldots, \mathbf{a}_j + \alpha\mathbf{a}_k, \ldots, \mathbf{a}_n) = \det(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)$ $(n > 1, j \neq k)$, *ou, plus généralement,* $\det(\mathbf{a}_1, \ldots, \mathbf{a}_j + \sum_{k:k\neq j} \alpha_k\mathbf{a}_k, \ldots, \mathbf{a}_n) = \det(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)$.

(b) $\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \neq 0$ *si et seulement si les vecteurs-colonnes* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ *sont linéairement indépendants.*

En d'autres termes, le déterminant reste inchangé lorsqu'on additionne à l'une de ses colonnes un multiple d'une autre colonne ou une combinaison linéaire des autres colonnes; en outre, vu la proposition 4.2.4, le déterminant d'une matrice est non nul si et seulement si cette matrice est inversible (ou nul si et seulement si cette matrice n'est pas inversible).

5.1.7 Déterminant d'une matrice élémentaire

Le déterminant d'une matrice élémentaire se calcule par application répétée de (5.1) ($|\mathbf{I}_i^a|$ et $|\mathbf{I}_{il}^a|$ avec $i < l$ sont des cas particuliers de (5.4)). Il peut également être déduit de (d) du théorème 5.1.5 et (a) ou (b) du même théorème, ou (a) de la proposition 5.1.6. Le résultat est le suivant:

$$|\mathbf{I}_{il}| = -1, \quad |\mathbf{I}_{il}^a| = 1, \quad |\mathbf{I}_i^a| = a. \tag{5.6}$$

Vu l'effet de la multiplication par une matrice élémentaire (cf. 4.4.2), il s'ensuit, grâce à (a) et (b) du théorème 5.1.5 et à (a) de la proposition 5.1.6, que

$$|\mathbf{MB}| = |\mathbf{M}||\mathbf{B}| \tag{5.7}$$

pour toute matrice élémentaire $\mathbf{M}$ et toute matrice carrée $\mathbf{B}$ du même ordre que $\mathbf{M}$.

De (5.6) et des relations (a), (b) et (c) de 4.4.3, il résulte en outre que

$$|{}^t\mathbf{M}| = |\mathbf{M}| \tag{5.8}$$

pour toute matrice élémentaire $\mathbf{M}$.

Les propriétés (5.7) et (5.8) serviront à démontrer leur généralisation que voici:

5.1.8 Proposition. Autres propriétés du déterminant d'ordre n

Si $\mathbf{A}$ *et* $\mathbf{B}$ *sont des matrices carrées d'ordre n, alors*

(a) $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$;
(b) $|{}^t\mathbf{A}| = |\mathbf{A}|$.

En d'autres termes, le déterminant du produit de deux matrices carrées du même ordre est égal au produit de leurs déterminants et le déterminant de la transposée d'une matrice carrée est égal au déterminant de cette matrice.

Il résulte aussitôt de (a) que

$$|\mathbf{A}^k| = |\mathbf{A}|^k \tag{5.9}$$

pour tout entier positif k.

5.1.9 Déterminant comme fonction des lignes

Le déterminant peut être conçu comme fonction de la famille des lignes d'une matrice. Grâce à (b) de la proposition 5.1.8, cette fonction jouit des propriétés énoncées dans le théorème 5.1.5 et dans la proposition 5.1.6, à savoir: elle est une forme n-linéaire alternée, elle attribue la valeur 1 à la famille des lignes de la matrice-unité, sa valeur ne change pas lorsqu'on additionne à une ligne un multiple d'une autre ligne et cette valeur est non nulle si et seulement si les lignes sont linéairement indépendantes.

5.1.10 Déterminant de l'inverse d'une matrice

Soit $\mathbf{A}$ une matrice inversible. D'après (d) du théorème 5.1.5 et (a) de la proposition 5.1.8,

$$1 = |\mathbf{I}| = |\mathbf{AA}^{-1}| = |\mathbf{A}||\mathbf{A}^{-1}|,$$

d'où nous déduisons que

$$|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}. \tag{5.10}$$

Il en découle aussitôt que dans le cas où $\mathbf{A}$ est inversible, la relation (5.9) est vraie pour tout entier k.

5.1.11 Proposition. Rang d'une matrice par les déterminants

Le rang d'une matrice $\mathbf{A}$ *(non nécessairement carrée) est le plus grand entier* k *tel que l'on puisse extraire de* $\mathbf{A}$ *une sous-matrice carrée d'ordre* k *et de déterminant non nul.*

DÉMONSTRATION

Soit r ce plus grand entier et m le nombre de lignes de $\mathbf{A} = (a_{ij})$. Désignons par $j_1, j_2, ..., j_r$ les indices de colonne d'une sous-matrice carrée de $\mathbf{A}$ de déterminant non nul. D'après (b) de la proposition 5.1.6, les colonnes et donc les lignes de cette sous-matrice sont linéairement indépendantes. Il s'ensuit que le rang de la matrice

$$\begin{pmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_r} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_r} \\ & \cdots\cdots & & \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_r} \end{pmatrix}$$

est r, car r de ses lignes sont linéairement indépendantes. Par conséquent, $\mathrm{rg}\mathbf{A} \geqslant r$. D'autre part, r' étant le rang de $\mathbf{A}$, nous pouvons extraire de $\mathbf{A}$ r' colonnes linéairement indépendantes formant une sous-matrice $\mathbf{A}'$ de rang r'. En prenant alors r' lignes linéairement indépendantes de $\mathbf{A}'$, nous constituons une sous-matrice carrée de $\mathbf{A}$ d'ordre r' et de rang r'. D'après (b) de la proposition 5.1.6, le déterminant de cette sous-matrice est non nul. Par conséquent, $r' = \mathrm{rg}\mathbf{A} \leqslant r$. Nous en concluons que $\mathrm{rg}\mathbf{A} = r$, ce qu'il fallait démontrer.

5.1.12 Volume d'un parallélépipède

Soit $\mathscr{E}$ un espace affine euclidien de dimension finie non nulle n. On appelle *parallélépipède* construit sur les points $P_0, P_1, ..., P_n$ de $\mathscr{E}$ le sous-ensemble de $\mathscr{E}$

$$\{P\colon \overrightarrow{P_0P} = \sum_{j=1}^{n} \alpha_j \overrightarrow{P_0P_j},\ 0 \leqslant \alpha_j \leqslant 1, j = 1, 2, ..., n\}. \tag{5.11}$$

Munissons $\mathscr{E}$ d'un repère orthonormal et désignons par $\mathbf{a}_j$ le vecteur-colonne de $\mathbb{R}^n$ formé des composantes du vecteur $\overrightarrow{P_0P_j}$. On appelle *volume* du parallélépipède (5.11), et on note $\mathrm{vol}(P_0, P_1, ..., P_n)$, la valeur absolue du déterminant des vecteurs-colonnes $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$:

$$\mathrm{vol}(P_0, P_1, ..., P_n) = |\det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)|. \tag{5.12}$$

Dans le cas où $n = 1$, un parallélépipède est un intervalle et son volume est la longueur.

Dans le cas où $n = 2$, un parallélépipède est appelé *parallélogramme* et son volume *aire*. D'après (5.12)

$$\mathrm{aire}(P_0, P_1, P_2) = |\det(\mathbf{a}_1, \mathbf{a}_2)| = \left|\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\right| = |a_{11}a_{22} - a_{21}a_{12}|.$$

Montrons que le dernier membre est bien l'aire ordinaire du parallélogramme construit sur P_0, P_1, P_2. Le carré de celle-ci est égal à

$$\begin{aligned} \|\overrightarrow{P_0P_1}\|^2 \|\overrightarrow{P_0P_2}\|^2 \sin^2(P_1\hat{P}_0P_2) &= \|\mathbf{a}_1\|^2 \|\mathbf{a}_2\|^2 - \|\mathbf{a}_1\|^2 \|\mathbf{a}_2\|^2 \cos^2(P_1\hat{P}_0P_2) \\ &= \|\mathbf{a}_1\|^2 \|\mathbf{a}_2\|^2 - (\mathbf{a}_1 | \mathbf{a}_2)^2 = (a_{11}a_{22} - a_{21}a_{12})^2, \end{aligned}$$

ce qui démontre l'assertion si P_1 et P_2 sont distincts de P_0; autrement, elle est évidente.

Compte tenu de 2.7.14 et de 2.7.15, nous voyons également que

$$\mathrm{vol}(P_0, P_1, P_2, P_3) = |\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)| = \left|\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}\right|$$

représente le volume ordinaire du parallélépipède construit sur P_0, P_1, P_2, P_3.

Nous montrerons dans 7.1.12 que la définition de volume (5.12) est indépendante du choix du repère orthonormal. De ce fait, le volume d'un cube-unité, c'est-à-dire d'un parallélépipède tel que les vecteurs $\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, ..., \overrightarrow{P_0P_n}$ sont orthogonaux deux à deux et unitaires, est égal à 1, car dans le repère $(P_0; \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, ..., \overrightarrow{P_0P_n})$ ce volume s'écrit

$$\mathrm{vol}(P_0, P_1, ..., P_n) = |\det(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)|$$

et vaut donc 1, par la propriété (d) du théorème 5.1.5.

A titre d'exercice, le lecteur pourra écrire les propriétés du volume qui découlent de celles du déterminant et illustrer ces propriétés, à l'aide d'une figure, dans le cas de l'aire et du volume ordinaires.

5.2 DÉMONSTRATIONS DES PROPRIÉTÉS DES DÉTERMINANTS

5.2.1 Démonstration du théorème 5.1.5

Propriété (a). Cette propriété est évidente si $n = 2$. Supposons qu'elle soit vérifée par le déterminant d'ordre $n - 1$, n étant supérieur à 2. En vertu de (5.1), elle est alors vérifiée par le déterminant d'ordre n dans le cas où $1 < j < k$. Le cas où $j = 1$ et $k > 2$ se réduit à celui où $j = 1$ et $k = 2$ par un triple échange: d'abord de $\mathbf{a}_2$ et $\mathbf{a}_k$, ensuite de $\mathbf{a}_1$ et $\mathbf{a}_k$ et finalement de $\mathbf{a}_1$ et $\mathbf{a}_2$. Le premier et le troisième

échange transforment chacun la valeur du déterminant en sa valeur opposée, par ce que nous venons de constater. Le deuxième échange en fait autant, en vertu de ce que nous allons démontrer. Pour tout couple (i, l) tel que $i \neq l$, désignons par $\mathbf{A}_{il,12}$ la matrice carrée d'ordre $n-2$ déduite de $\mathbf{A}$ par suppression de la i-ième et la l-ième ligne, ainsi que des deux premières colonnes. D'après (5.1),

$$|\mathbf{A}_{i1}| = \sum_{l=1}^{i-1}(-1)^{l+1}a_{l2}|\mathbf{A}_{il,12}| + \sum_{l=i+1}^{n}(-1)^{l}a_{l2}|\mathbf{A}_{il,12}|,$$

où il est entendu que le second membre se réduit à la deuxième somme si $i = 1$ et à la première si $i = n$. Substituons ce second membre à $|\mathbf{A}_{i1}|$ dans (5.1). Il vient:

$$\begin{aligned} |\mathbf{A}| &= \sum_{i=2}^{n}\sum_{l=1}^{i-1}(-1)^{i+l}a_{i1}a_{l2}|\mathbf{A}_{il,12}| - \sum_{i=1}^{n-1}\sum_{l=i+1}^{n}(-1)^{i+l}a_{i1}a_{l2}|\mathbf{A}_{il,12}| \\ &= \sum_{(i,l)\in\Delta_1}(-1)^{i+l}a_{i1}a_{l2}|\mathbf{A}_{il,12}| - \sum_{(i,l)\in\Delta_2}(-1)^{i+l}a_{i1}a_{l2}|\mathbf{A}_{il,12}|, \end{aligned} \tag{5.13}$$

où $\Delta_1 = \{(i, l)\colon\ 2 \leqslant i \leqslant n,\ \ 1 \leqslant l \leqslant i-1\}$ et $\Delta_2 = \{(i, l)\colon\ 1 \leqslant i \leqslant l-1,\ 2 \leqslant l \leqslant n\}$. Si nous échangeons la première et la deuxième colonne de $\mathbf{A}$, les matrices $\mathbf{A}_{il,12}$ ne subissent aucune modification. En revanche, $a_{i1}a_{l2}$ change en $a_{l1}a_{i2}$, pour tout couple (i, l). Nous voyons ainsi que la deuxième somme du dernier membre de (5.13) devient la première et, inversement, la première devient la deuxième. L'échange de la première et la deuxième colonne a donc pour seul effet de multiplier le déterminant de $\mathbf{A}$ par -1, ce qu'il fallait démontrer.

Propriétés (b) *et* (c). Vu la propriété (a), nous pouvons admettre que $j = 1$. Les conclusions découlent alors directement de (5.1).

Propriété (d). Il suffit de poser $a = 1$ dans (5.3).

5.2.2 Démonstration de la proposition 5.1.6

Propriété (a). En vertu de (a) du théorème 5.1.5, nous pouvons admettre que $j = 1$. D'après (b) et (c) du même théorème,

$$\det(\mathbf{a}_1 + \alpha\mathbf{a}_k, \mathbf{a}_2, ..., \mathbf{a}_n) = \det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) + \alpha\det(\mathbf{a}_k, \mathbf{a}_2, ..., \mathbf{a}_n).$$

Puisque k est différent de $j = 1$ par hypothèse, le vecteur-colonne $\mathbf{a}_k$ apparaît deux fois dans le dernier déterminant. Celui-ci est donc nul, d'après ce qui a été remarqué dans 5.1.5. La formulation générale s'ensuit par itération.

Propriété (b). Supposons que les vecteurs-colonnes $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ soient linéairement dépendants. En vertu de la proposition 1.5.5, un de ces vecteurs, par exemple $\mathbf{a}_j$, est combinaison linéaire des autres:

$$\mathbf{a}_j = \sum_{k:k\neq j}\alpha_k\mathbf{a}_k.$$

Il s'ensuit, grâce à la propriété (a) déjà établie, que

$$\begin{aligned} \det(\mathbf{a}_1, ..., \mathbf{a}_j, ..., \mathbf{a}_n) &= \det(\mathbf{a}_1, ..., \mathbf{a}_j - \sum_{k:k\neq j}\alpha_k\mathbf{a}_k, ..., \mathbf{a}_n) \\ &= \det(\mathbf{a}_1, ..., \mathbf{0}, ..., \mathbf{a}_n) = 0. \end{aligned}$$

Inversement, supposons que les vecteurs-colonnes $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ soient linéairement indépendants. Selon la proposition 4.2.4, la matrice $\mathbf{A}$ de colonnes $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ est alors inversible. D'après (d) du théorème 5.1.5 et (a) de la proposition 5.1.8 (démontrée ci-dessous),

$$1 = |\mathbf{I}| = |\mathbf{A}\mathbf{A}^{-1}| = |\mathbf{A}||\mathbf{A}^{-1}|,$$

ce qui entraîne que $|\mathbf{A}| = \det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)$ est non nul.

5.2.3 Démonstration de la proposition 5.1.8

Propriété (a). Si $\mathbf{A}$ n'est pas inversible, le produit $\mathbf{AB}$ ne l'est pas non plus, selon (c) de la proposition 4.2.6. Or, nous venons d'établir que le déterminant d'une matrice non inversible (c'est-à-dire dont les colonnes sont linéairement dépendantes) est nul. L'égalité $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ est donc vraie dans ce cas. Si $\mathbf{A}$ est inversible, il existe, d'après 4.4.4, des matrices élémentaires $\mathbf{M}_1, \mathbf{M}_2, ..., \mathbf{M}_k$ telles que $\mathbf{A} = \mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k$. Compte tenu de (5.7), nous pouvons donc écrire

$$\begin{aligned} |\mathbf{AB}| &= |\mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k\mathbf{B}| = |\mathbf{M}_1||\mathbf{M}_2 ... \mathbf{M}_k\mathbf{B}| = ... \\ &= |\mathbf{M}_1||\mathbf{M}_2| ... |\mathbf{M}_k||\mathbf{B}| = |\mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k||\mathbf{B}| = |\mathbf{A}||\mathbf{B}|, \end{aligned}$$

ce qui établit la propriété (a).

Propriété (b). Si $\mathbf{A}$ n'est pas inversible, ${}^t\mathbf{A}$ ne l'est pas non plus, d'après (a) de 4.2.6, donc $|\mathbf{A}| = |{}^t\mathbf{A}| = 0$, d'après (b) de la proposition 5.1.6. Si $\mathbf{A}$ est inversible, $\mathbf{A} = \mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k$, où $\mathbf{M}_1, \mathbf{M}_2, ..., \mathbf{M}_k$ sont des matrices élémentaires. Dans ce cas, grâce à la relation (e) de 4.1.9, à la propriété (a) démontrée ci-dessus et à (5.8), nous pouvons écrire

$$\begin{aligned} |{}^t\mathbf{A}| &= |{}^t\mathbf{M}_k{}^t\mathbf{M}_{k-1} ... {}^t\mathbf{M}_1| = |{}^t\mathbf{M}_k||{}^t\mathbf{M}_{k-1}| ... |{}^t\mathbf{M}_1| \\ &= |\mathbf{M}_1||\mathbf{M}_2| ... |\mathbf{M}_k| = |\mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k| = |\mathbf{A}|, \end{aligned}$$

ce qui établit la propriété (b).

5.2.4 Caractérisation du déterminant d'ordre *n*

Le déterminant d'ordre n est la seule fonction réelle définie dans l'ensemble des matrices carrées d'ordre n vérifiant les propriétés (a)–(d) énoncées dans le théorème 5.1.5. En effet, nous avons démontré qu'une telle fonction, tout comme le déterminant d'ordre n, attribue à $\mathbf{A}$ la valeur 0 si $\mathbf{A}$ est une matrice non inversible et la valeur $|\mathbf{M}_1||\mathbf{M}_2| ... |\mathbf{M}_k|$ si $\mathbf{A}$ est un produit de matrices élémentaires $\mathbf{M}_1, \mathbf{M}_2, ..., \mathbf{M}_k$. Toute matrice carrée étant de l'une de ces deux sortes, l'assertion est démontrée.

5.3 DÉVELOPPEMENTS, FORMULE DE CRAMER

5.3.1 Développement d'un déterminant suivant une colonne

Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre $n > 1$ et $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ ses colonnes. Fixons $j > 1$ et désignons par $\mathbf{A}' = (a'_{ij})$ la matrice obtenue à partir de $\mathbf{A}$ en échangeant successivement $\mathbf{a}_j$ et $\mathbf{a}_{j-1}$, $\mathbf{a}_j$ et $\mathbf{a}_{j-2}$, ..., $\mathbf{a}_j$ et $\mathbf{a}_1$. Les colonnes de $\mathbf{A}'$ (écrites dans l'ordre) étant $\mathbf{a}_j, \mathbf{a}_1, ..., \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, ..., \mathbf{a}_n$, il est clair que, pour tout indice de ligne i,

$$a'_{i1} = a_{ij} \quad \text{et} \quad \mathbf{A}'_{i1} = \mathbf{A}_{ij}.$$

D'après (5.1), le déterminant de $\mathbf{A}'$ s'écrit donc sous la forme

$$|\mathbf{A}'| = \sum_{i=1}^{n} (-1)^{i+1} a_{ij} |\mathbf{A}_{ij}|. \tag{5.14}$$

D'autre part, puisque $\mathbf{A}'$ est déduite de $\mathbf{A}$ par $j-1$ échanges de colonnes, la propriété (a) du théorème 5.1.5 entraîne

$$|\mathbf{A}'| = (-1)^{j-1} |\mathbf{A}|. \tag{5.15}$$

En comparant (5.14) à (5.15), il résulte que

$$|\mathbf{A}| = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |\mathbf{A}_{ij}|, \quad j = 1, 2, ..., n. \tag{5.16}$$

On appelle le second membre de (5.16) *développement du déterminant de* $\mathbf{A}$ *suivant la j-ième colonne*. On remarquera que (5.1) est un cas particulier de (5.16), à savoir le cas où $j = 1$.

5.3.2 Développement d'un déterminant suivant une ligne

Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre $n > 1$. D'après (5.16), le développement de la matrice ${}^t\mathbf{A} = (a'_{ji})$ suivant la i-ième colonne s'écrit

$$|{}^t\mathbf{A}| = \sum_{j=1}^{n} (-1)^{i+j} a'_{ji} |({}^t\mathbf{A})_{ji}|. \tag{5.17}$$

Comme

$$a'_{ji} = a_{ij} \quad \text{et} \quad ({}^t\mathbf{A})_{ji} = {}^t(\mathbf{A}_{ij}),$$

il s'ensuit, vu la propriété (b) de la proposition 5.1.8, que

$$|\mathbf{A}| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |\mathbf{A}_{ij}|, \quad i = 1, 2, ..., n. \tag{5.18}$$

On appelle le second membre de (5.18) *développement du déterminant de* $\mathbf{A}$ *suivant la i-ième ligne*.

5.3.3 Mineurs, cofacteurs

On appelle $|\mathbf{A}_{ij}|$ *mineur* d'indices i, j et $(-1)^{i+j} |\mathbf{A}_{ij}|$ *cofacteur* de a_{ij}. Le signe $(-1)^{i+j}$ varie selon le tableau suivant:

$$\begin{pmatrix} + & - & + & \cdot & \cdot & \cdot \\ - & + & - & \cdot & \cdot & \cdot \\ + & - & + & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & + & - \\ \cdot & \cdot & \cdot & \cdot & - & + \end{pmatrix}.$$

Ainsi, d'après (5.16) et (5.18), le déterminant d'une matrice carrée est égal à la somme des termes d'une colonne ou d'une ligne multipliés par leurs cofacteurs respectifs.

5.3.4 Exemple d'application

Les développements s'emploient avantageusement pour calculer le déterminant d'une matrice dont plusieurs termes sont nuls. Dans l'exemple suivant, la flèche indique la ligne ou la colonne suivant laquelle le déterminant est développé:

$$\begin{vmatrix} 1 & 0 & -3 & 0 & 7 & 2 \\ 11 & -3 & 2 & 1 & 10 & -2 \\ 0 & 0 & 1 & 0 & 6 & 0 \\ 5 & 1 & 4 & 0 & -9 & 6 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ -3 & 0 & -10 & 0 & 4 & 1 \end{vmatrix}_{\substack{\\ \qquad\qquad\qquad \uparrow}} = \begin{vmatrix} 1 & 0 & -3 & 7 & 2 \\ 0 & 0 & 1 & 6 & 0 \\ 5 & 1 & 4 & -9 & 6 \\ 0 & 0 & 0 & 2 & 0 \\ -3 & 0 & -10 & 4 & 1 \end{vmatrix}_{\substack{\\ \qquad \uparrow}}$$

$$= -\begin{vmatrix} 1 & -3 & 7 & 2 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 2 & 0 \\ -3 & -10 & 4 & 1 \end{vmatrix} \begin{matrix} \\ \\ \leftarrow \\ \\ \end{matrix} = -2 \begin{vmatrix} 1 & -3 & 2 \\ 0 & 1 & 0 \\ -3 & -10 & 1 \end{vmatrix} \begin{matrix} \\ \leftarrow \\ \\ \end{matrix}$$

$$= -2 \begin{vmatrix} 1 & 2 \\ -3 & 1 \end{vmatrix} = -2 \cdot 7 = -14.$$

5.3.5 Formule de Cramer

Considérons un système de n équations à n inconnues, de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$. Supposons que le déterminant de $\mathbf{A}$ soit non nul. D'après la proposition 3.5.9, ce système admet une solution unique $\mathbf{x}_0 = (x_j^0)$. Nous allons exposer une méthode de calcul de $\mathbf{x}_0$ s'appuyant sur les propriétés des déterminants. Désignons par $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ les colonnes de $\mathbf{A}$. Par hypothèse,

$$x_1^0\mathbf{a}_1 + x_2^0\mathbf{a}_2 + ... + x_n^0\mathbf{a}_n = \mathbf{b}.$$

Substituons le premier membre de cette égalité à $\mathbf{b}$ dans le déterminant

$$\det(\mathbf{a}_1, ..., \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, ..., \mathbf{a}_n).$$

D'après (a) de la proposition 5.1.6, ce déterminant est égal à

$$\det(\mathbf{a}_1, ..., \mathbf{a}_{j-1}, x_j^0\mathbf{a}_j, \mathbf{a}_{j+1}, ..., \mathbf{a}_n) = x_j^0\det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n).$$

Il s'ensuit que

$$x_j^0 = \frac{1}{|\mathbf{A}|}\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ & & \cdots & & \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}, \quad j = 1, 2, ..., n. \tag{5.19}$$

$$\uparrow \quad j\text{-ième}$$

Cette relation est connue sous le nom de *formule de Cramer* pour la résolution d'un système linéaire.

Résolvons, par exemple, le système linéaire

$$\begin{aligned} 3x_1 + 2x_2 + 6x_3 &= 5 \\ 2x_1 + 3x_2 + 3x_3 &= 2 \\ 2x_1 - 2x_2 - 4x_3 &= 1. \end{aligned}$$

Selon (5.19),

$$x_1 = \frac{\begin{vmatrix} 5 & 2 & 6 \\ 2 & 3 & 3 \\ 1 & -2 & -4 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 6 \\ 2 & 3 & 3 \\ 2 & -2 & -4 \end{vmatrix}} = 1, \quad x_2 = \frac{\begin{vmatrix} 3 & 5 & 6 \\ 2 & 2 & 3 \\ 2 & 1 & -4 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 6 \\ 2 & 3 & 3 \\ 2 & -2 & -4 \end{vmatrix}} = -\tfrac{1}{2}, \quad x_3 = \frac{\begin{vmatrix} 3 & 2 & 5 \\ 2 & 3 & 2 \\ 2 & -2 & 1 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 6 \\ 2 & 3 & 3 \\ 2 & -2 & -4 \end{vmatrix}} = \tfrac{1}{2}.$$

5.3.6 Calcul de la matrice inverse par la méthode des cofacteurs

Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre n et de déterminant non nul. D'après (b) de la proposition 5.1.6 et la proposition 4.2.4, $\mathbf{A}$ est inversible. Les colonnes de la matrice inverse $\mathbf{A}^{-1}$ sont les solutions des n équations matricielles (4.13):

$$\mathbf{A}\begin{pmatrix} x_{1j} \\ x_{2j} \\ \cdot \\ \cdot \\ \cdot \\ x_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 1 \\ \cdot \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow j\text{-ième} \\ \\ \\ \end{matrix}, \qquad j = 1, 2, ..., n.$$

Ces équations sont l'expression matricielle de n systèmes linéaires que nous résolvons à l'aide de la formule de Cramer (5.19). La solution unique du j-ième système s'écrit

$$x_{ij} = \frac{1}{|\mathbf{A}|}\begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ & & \cdots & & \\ a_{j1} & \cdots & 1 & \cdots & a_{jn} \\ & & \cdots & & \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix}, \quad i = 1, 2, ..., n,$$

$$\uparrow \quad i\text{-ième}$$

ou encore, après développement du dernier déterminant suivant la i-ième colonne,

$$x_{ij} = \frac{(-1)^{i+j}|\mathbf{A}_{ji}|}{|\mathbf{A}|} = \frac{\text{cof}\,a_{ji}}{|\mathbf{A}|}, \quad i = 1, 2, ..., n, \tag{5.20}$$

où $\text{cof}\,a_{ji}$ désigne le cofacteur de a_{ji}. Posons $\tilde{\mathbf{A}} = (\tilde{a}_{ij})$, où $\tilde{a}_{ij} = \text{cof}\,a_{ji}$. La matrice inverse de $\mathbf{A}$ s'écrit alors sous la forme

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\tilde{\mathbf{A}}. \qquad A^{-1} = \frac{1}{|A|}\,\text{adj}\,A \tag{5.21}$$

En d'autres termes, $\mathbf{A}^{-1}$ est égale à la matrice des cofacteurs transposée divisée par le déterminant de $\mathbf{A}$.

On notera le cas particulier où $n = 2$:

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \tag{5.22}$$

Calculons, par exemple, la matrice inverse de .

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & 2 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix}.$$

En développant le déterminant de **A** suivant la troisième ligne, nous voyons aisément que $|\mathbf{A}| = -2$. D'autre part,

$$\operatorname{cof} a_{11} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1, \quad \operatorname{cof} a_{12} = -\begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = -1,$$

$$\operatorname{cof} a_{13} = \begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} = 1, \quad \operatorname{cof} a_{21} = -\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} = 2, \ldots,$$

$$\tilde{\mathbf{A}} = {}^t\begin{pmatrix} -1 & -1 & 1 \\ 2 & 2 & -4 \\ -1 & -3 & 3 \end{pmatrix}.$$

D'après (5.21), nous concluons que

$$\mathbf{A}^{-1} = -\tfrac{1}{2}\begin{pmatrix} -1 & 2 & -1 \\ -1 & 2 & -3 \\ 1 & -4 & 3 \end{pmatrix}.$$

5.4 EXEMPLES ET REMARQUES DIVERSES

5.4.1 Quelques exemples de calcul d'un déterminant

Le calcul d'un déterminant est d'ordinaire précédé d'un travail de simplification fait principalement à l'aide de (a) de la proposition 5.1.6 ou de la propriété correspondante relative aux lignes.

(1) Pour calculer le déterminant

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{vmatrix},$$

on peut commencer par additionner la première colonne à la troisième et à la quatrième et appliquer (b) du théorème 5.1.5 à ces deux colonnes:

$$\begin{vmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 2 & 0 \\ 1 & -1 & 2 & 2 \end{vmatrix} = 4\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{vmatrix}.$$

Ensuite, il convient de soustraire de la première ligne deux fois la deuxième et une fois la quatrième:

$$4\begin{vmatrix} -2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{vmatrix} = 4\cdot(-2) = -8.$$

(2) Pour calculer le déterminant

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix},$$

on peut soustraire la deuxième ligne de la première et la troisième de la deuxième:

$$\begin{vmatrix} 0 & a-b & a^2-b^2 \\ 0 & b-c & b^2-c^2 \\ 1 & c & c^2 \end{vmatrix} = (a-b)(b-c)\begin{vmatrix} 1 & a+b \\ 1 & b+c \end{vmatrix} = (a-b)(b-c)(c-a).$$

(3) On notera, à titre d'exemple, que

$$\begin{vmatrix} 2 & -3 & 1 \\ 2 & 3 & -2 \\ 4 & 0 & -1 \end{vmatrix} = 0,$$ car la troisième ligne est la somme des deux premières.

(4) Le déterminant

$$\begin{vmatrix} 1+a & 1 & 1 & 1 \\ 1 & 1+b & 1 & 1 \\ 1 & 1 & 1+c & 1 \\ 1 & 1 & 1 & 1+d \end{vmatrix}$$

peut être calculé en appliquant (c) du théorème 5.1.5 et (a) de la proposition 5.1.6. A cet effet, posons

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ b \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ c \\ 0 \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ d \end{pmatrix}.$$

Alors le déterminant s'écrit:

$$\begin{aligned}
&\det(\mathbf{v}+\mathbf{a}, \mathbf{v}+\mathbf{b}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) \\
&= \det(\mathbf{v}, \mathbf{v}+\mathbf{b}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) + \det(\mathbf{a}, \mathbf{v}+\mathbf{b}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) \\
&= \det(\mathbf{v}, \mathbf{b}, \mathbf{c}, \mathbf{d}) + \det(\mathbf{a}, \mathbf{v}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) \\
&= bcd + \det(\mathbf{a}, \mathbf{v}, \mathbf{c}, \mathbf{d}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{v}+\mathbf{d}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{v}+\mathbf{d}) \\
&= bcd + acd + \det(\mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{d}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{v}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) \\
&= bcd + acd + abd + abc + abcd.
\end{aligned}$$

(5) Nous vous proposons de calculer le déterminant de la matrice carrée d'ordre n

$$\mathbf{A}_n = \begin{bmatrix} 2 & 1 & 0 & 0 & . & . & . & 0 \\ 1 & 2 & 1 & 0 & . & . & . & 0 \\ 0 & 1 & 2 & 1 & . & . & . & 0 \\ & & & . & . & . & . & . & . \\ 0 & . & . & . & . & 1 & 2 & 1 \\ 0 & . & . & . & . & 0 & 1 & 2 \end{bmatrix}.$$

Supposons $n > 2$ et développons $|\mathbf{A}_n|$ suivant la première ligne:

$$|\mathbf{A}_n| = 2|\mathbf{A}_{n-1}| - \begin{vmatrix} 1 & 1 \; 0 \; . \; . \; . \; 0 \\ 0 & \\ . & \\ . & \mathbf{A}_{n-2} \\ . & \\ 0 & \end{vmatrix} = 2|\mathbf{A}_{n-1}| - |\mathbf{A}_{n-2}|.$$

Vu que $|\mathbf{A}_1| = 2$ et $|\mathbf{A}_2| = 3$, cette formule nous permet de calculer $|\mathbf{A}_n|$ par récurrence. Le résultat est $|\mathbf{A}_n| = n + 1$.

5.4.2 Déterminant d'une matrice partagée en sous-matrices

Commençons par observer que, pour toute matrice carrée $\mathbf{A}$ d'ordre m et toute matrice $\mathbf{C}$ de type $m \times n$

$$\left|\begin{array}{c:c} \mathbf{I}_n & \mathbf{O} \\ \hdashline \mathbf{C} & \mathbf{A} \end{array}\right| = \left|\begin{array}{c:c} \mathbf{A} & \mathbf{C} \\ \hdashline \mathbf{O} & \mathbf{I}_n \end{array}\right| = |\mathbf{A}|. \tag{5.23}$$

Pour s'en rendre compte, il suffit de développer le premier (deuxième) déterminant n fois suivant la première (dernière) ligne.

Soit maintenant $\mathbf{A}$ et $\mathbf{B}$ des matrices carrées respectivement d'ordre m et n et $\mathbf{C}$ une matrice de type $m \times n$. En vertu de (4.7),

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{I}_n \end{bmatrix}.$$

Il s'ensuit, grâce à (a) de la proposition 5.1.8 et à (5.23), que

$$\left|\begin{array}{c:c} \mathbf{A} & \mathbf{C} \\ \hdashline \mathbf{O} & \mathbf{B} \end{array}\right| = |\mathbf{A}||\mathbf{B}|. \tag{5.24}$$

On notera cependant que si $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ et $\mathbf{D}$ sont des matrices carrées du même ordre, en général

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{D} & \mathbf{B} \end{vmatrix} \neq |\mathbf{A}||\mathbf{B}| - |\mathbf{D}||\mathbf{C}|.$$

Pour s'en rendre compte, il suffit de prendre $\mathbf{A} = \mathbf{B} = a\mathbf{I}_2$ et $\mathbf{C} = \mathbf{D} = c\mathbf{I}_2$, $a \neq c$ (cf. exercice 5.5.11).

5.4.3 Retour au volume d'un parallélépipède

Considérons à nouveau le parallélépipède défini par (5.11). Munissons l'espace affine euclidien $\mathscr{E}$ d'un repère orthonormal et désignons par $x_{j1}, x_{j2}, \ldots, x_{jn}$ les coordonnées de P_j, où $j = 0, 1, \ldots, n$. Le volume $\operatorname{vol}(P_0, P_1, \ldots, P_n)$ de ce parallélépipède est alors la valeur absolue du déterminant

$$\begin{vmatrix} 1 & x_{01} & x_{02} & . & . & . & x_{0n} \\ 1 & x_{11} & x_{12} & . & . & . & x_{1n} \\ & & . & . & . & . & . & . \\ 1 & x_{n1} & x_{n2} & . & . & . & x_{nn} \end{vmatrix}. \tag{5.25}$$

Pour établir cela, il suffit de soustraire la première ligne de toutes les autres et d'appliquer (5.1). Compte tenu de (b) de la proposition 5.1.8, le résultat de ces opérations est le déterminant $\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ dont la valeur absolue est, selon la définition (5.12), le volume du parallélépipède.

5.4.4 Equation cartésienne d'un hyperplan

Soit $\mathscr{E}$ un espace affine de dimension finie non nulle n et $P_1, P_2, \ldots, P_n$ n points de $\mathscr{E}$ engendrant un hyperplan. Supposons que $\mathscr{E}$ soit muni d'un repère et désignons par $x_{j1}, x_{j2}, \ldots, x_{jn}$ les coordonnées de P_j, où $j = 1, 2, \ldots, n$. L'équation cartésienne de cet hyperplan peut alors s'écrire sous la forme

$$\begin{vmatrix} 1 & x_1 & x_2 & . & . & . & x_n \\ 1 & x_{11} & x_{12} & . & . & . & x_{1n} \\ & & . & . & . & . & . & . \\ 1 & x_{n1} & x_{n2} & . & . & . & x_{nn} \end{vmatrix} = 0. \tag{5.26}$$

En effet, en développant le déterminant suivant la première ligne, nous obtenons une équation de la forme (1.21). En outre, chaque point P_j appartient à l'hyperplan défini par cette équation, car en substituant les coordonnées $x_{j1}, x_{j2}, \ldots, x_{jn}$ de P_j aux inconnues $x_1, x_2, \ldots, x_n$, nous voyons que le déterminant a deux lignes égales et donc qu'il est nul.

5.5 EXERCICES

5.5.1 Montrer que

$$\begin{vmatrix} 1 & 1 & 3 & 0 & 2 \\ 3 & 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & 0 & 2 \\ 4 & -2 & 3 & 1 & 0 \\ 5 & 1 & 0 & 0 & 6 \end{vmatrix} = 60.$$

5.5.2 Calculer le déterminant de la matrice

$$\mathbf{A} = \begin{pmatrix} a & 1 & -1 & 1 \\ -1 & a & -1 & -1 \\ 1 & 1 & a & -1 \\ -1 & 1 & 1 & a \end{pmatrix}$$

(a) d'une manière directe,
(b) en effectuant, au préalable, le produit $\mathbf{A}^t\mathbf{A}$.

5.5.3 A l'aide du théorème 5.1.5 et de la proposition 5.1.6, montrer que

$$\begin{vmatrix} \alpha a_1+b_1 & \alpha b_1+c_1 & \alpha c_1+a_1 \\ \alpha a_2+b_2 & \alpha b_2+c_2 & \alpha c_2+a_2 \\ \alpha a_3+b_3 & \alpha b_3+c_3 & \alpha c_3+a_3 \end{vmatrix} = (\alpha^3+1)\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

5.5.4 Quel est le degré du polynôme p défini par

$$p(t) \equiv \begin{vmatrix} a_{11}+t & a_{12}+t & \dots & a_{1n}+t \\ a_{21}+t & a_{22}+t & \dots & a_{2n}+t \\ & \cdots\cdots & & \\ a_{n1}+t & a_{n2}+t & \dots & a_{nn}+t \end{vmatrix} ?$$

5.5.5 Soit $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ une famille de vecteurs d'un espace vectoriel euclidien. Démontrer que cette famille est liée si et seulement si son *déterminant de Gram*

$$\begin{vmatrix} (\mathbf{x}_1 \mid \mathbf{x}_1) & \dots & (\mathbf{x}_1 \mid \mathbf{x}_k) \\ & \cdots\cdots & \\ (\mathbf{x}_k \mid \mathbf{x}_1) & \dots & (\mathbf{x}_k \mid \mathbf{x}_k) \end{vmatrix}$$

est nul. (Utiliser la proposition 1.5.5 et les propriétés du déterminant.)

5.5.6 Montrer que pour toute famille de nombres $(a_1, a_2, ..., a_n)$ $(n > 1)$,

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ & & \cdots\cdots & & \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix} = \prod_{i>j} (a_i - a_j).$$

Ce déterminant est appelé *déterminant de Vandermonde*.

5.5.7 Résoudre le système de l'exercice 3.6.10 au moyen de la formule de Cramer.

5.5.8 Calculer la matrice inverse de

$$\begin{pmatrix} 1 & 2 & -2 \\ 1 & 3 & -1 \\ 2 & 1 & -5 \end{pmatrix}$$

par la méthode des cofacteurs.

5.5.9 Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre $n > 1$.

(a) Montrer que

$\tilde{\mathbf{A}}\mathbf{A} = |\mathbf{A}|\,\mathbf{I}_n.$

(Appliquer (5.16) à la matrice déduite de $\mathbf{A}$ en remplaçant sa j-ième colonne par sa k-ième.)

(b) Dans l'hypothèse que $|\mathbf{A}| \neq 0$, en déduire la formule d'inversion (5.21), ainsi que la relation

$|\tilde{\mathbf{A}}| = |\mathbf{A}|^{n-1}.$

(Utiliser (a) de la proposition 5.1.8 et (5.3).)

5.5.10 Démontrer, par récurrence sur le nombre n de lignes, que

$$\begin{vmatrix} 2\cos\theta & 1 & & & & \\ 1 & 2\cos\theta & 1 & & 0 & \\ & 1 & 2\cos\theta & 1 & & \\ & & \cdot & \cdot & \cdot & \\ & 0 & & \cdot & \cdot & 1 \\ & & & & 1 & 2\cos\theta \end{vmatrix} = \begin{cases} \dfrac{\sin((n+1)\theta)}{\sin\theta} & \text{si } \theta \neq k\pi, \\ n+1 & \text{si } \theta = 2k\pi, \\ (-1)^n(n+1) & \text{si } \theta = (2k+1)\pi. \end{cases}$$

5.5.11 Soit **A** une matrice inversible d'ordre m, **B** une matrice carrée d'ordre n, **C** une matrice de type $m \times n$ et **D** une matrice de type $n \times m$.

(a) Montrer que

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{D} & \mathbf{B} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{B} - \mathbf{D}\mathbf{A}^{-1}\mathbf{C}|.$$

(Vérifier que

$$\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{D} & \mathbf{B} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{D}\mathbf{A}^{-1} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} - \mathbf{D}\mathbf{A}^{-1}\mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \mathbf{A}^{-1}\mathbf{C} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix}$$

et appliquer (5.24).)

(b) En déduire que si **A** et **B** sont du même ordre et **A** et **D** commutent

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{D} & \mathbf{B} \end{vmatrix} = |\mathbf{AB} - \mathbf{DC}|.$$

5.5.12 Soit $P_1(a_1, a_2)$, $P_2(b_1, b_2)$ et $P_3(c_1, c_2)$ trois points non alignés d'un plan affine euclidien, muni d'un repère orthonormal. Montrer que

$$\begin{vmatrix} 1 & x_1^2+x_2^2 & x_1 & x_2 \\ 1 & a_1^2+a_2^2 & a_1 & a_2 \\ 1 & b_1^2+b_2^2 & b_1 & b_2 \\ 1 & c_1^2+c_2^2 & c_1 & c_2 \end{vmatrix} = 0$$

est l'équation du cercle passant par P_1, P_2 et P_3.

5.5.13 Dans un espace affine euclidien $\mathscr{E}$ de dimension finie n supérieure à 2 et muni d'un repère orthonormal, on considère un hyperplan $\mathscr{S}$ dont le j-ième vecteur directeur a pour composantes les termes de la j-ième colonne de la matrice

$$\mathbf{V} = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1,n-1} \\ v_{21} & v_{22} & \cdots & v_{2,n-1} \\ & & \cdots\cdots & \\ v_{n1} & v_{n2} & \cdots & v_{n,n-1} \end{pmatrix}.$$

(a) Soit $\mathbf{V}_i$ ($i = 1, 2, \ldots, n$) la matrice obtenue en supprimant la i-ième ligne de **V**. Montrer que le vecteur **n** de composantes

$$v_i = (-1)^{i+n}|\mathbf{V}_i|, \quad i = 1, 2, \ldots, n,$$

est normal à $\mathscr{S}$. (Appliquer (a) de l'exercice 5.5.9 à la matrice $\mathbf{A} = (\mathbf{V} \,\vdots\, \mathbf{a})$, où **a** est un vecteur-colonne quelconque de $\mathbb{R}^n$.)

(b) Interpréter ce résultat dans le cas où $n = 3$.

Index

Les nombres en caractère romain renvoient aux pages du tome 1
Les nombres en caractère italique renvoient aux pages du tome 2

Bibliographie

M. Berger, *Géométrie*, Cedic/Fernand Nathan, Paris, 1977.

A. Calame, *Mathématiques modernes*, Editions du Griffon, Neuchâtel, 1968.

L. Chambadal, J.L. Ovaert, *Algèbre linéaire et algèbre tensorielle*, Dunod, Paris, 1968.

J. Dieudonné, *Algèbre linéaire et géométrie élémentaire*, Hermann, Paris, 1964.

J. Frenkel, *Géométrie pour l'élève-professeur*, Hermann, Paris, 1973.

F.R. Gantmacher, *The theory of matrices*, Chelsea Publishing Company, New York, 1977.

R.J. Goult, *Applied linear algebra*, Ellis Horwood Ltd., Chichester, 1978.

W. Greub, *Linear algebra*, Springer-Verlag, New York, 1981.

P.R. Halmos, *Finite-dimensional vector spaces*, Van Nostrand, New York, 1958.

K. Hoffman, R. Kunze, *Linear algebra*, Prentice-Hall, Englewood Cliffs, New Jersey, 1961.

M. Jeger, B. Eckmann, *Einführung in die vektorielle Geometrie und lineare Algebra*, Birkhäuser Verlag, Basel, 1967.

H.-J. Kowalsky, *Lineare Algebra*, W. de Gruyter, Berlin, 1970.

P. Lancaster, M. Tismenetsky, *The theory of matrices*, Academic Press, Orlando, Florida, 1985.

S. Lang, *Algèbre linéaire*, Inter Editions, Paris, 1976.

H. Liebeck, *Algebra for scientists & engineers*, J. Wiley & Sons, London, New York, 1969.

S. Mac Lane, G. Birkhoff, *Algèbre*, Gauthier-Villars, Paris, 1971.

B. Noble, *Applied linear algebra*, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

E. Ramis, C. Deschamps, J. Odoux, *Cours de mathématiques spéciales*, Masson, Paris, 1977.

E. Snapper, R.J. Troyer, *Metric affine geometry*, Academic Press, New York, London, 1971.

E. Sperner, *Einführung in die Analytische Geometrie und Algebra*, Vandenhoeck & Ruprecht, Göttingen, 1957.

C. Tisseron, *Géométries affine, projective et euclidienne*, Hermann, Paris, 1983.